

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 13:46:45 ; Search time 11231 Seconds
3.98i Million cell updates/sec

Title: US-10-664-775-4

Perfect score: 2279

Sequence: 1 gatcactccctctagtgaag.....ttgtaattcttagtgctgat 2279

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 45 seqs, 9809608 residues

Total number of hits satisfying chosen parameters: 90

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rge4.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
	1 840.902	36.9	183326	1	AC124438
	2 840.902	36.9	287253	1	AC147502
C	3 780.401	34.2	230813	1	AC114097
	4 778.202	34.1	192016	1	AC102426
C	5 778.202	34.1	227550	1	AC102686
	6 768.303	33.7	217398	1	AC102519
	7 761.303	33.4	207687	1	AL824706
	8 757.703	33.2	158685	1	AC102201
	9 754.401	33.1	206312	1	AC102739
C	10 752.402	33.0	243010	1	AC120175
	11 752.302	33.0	236604	1	AC099619
	12 752.002	33.0	177414	1	AC124411
	13 729.201	32.0	184474	1	AL672047
	14 729.201	32.0	204031	1	AC093354
C	15 725.001	31.8	211131	1	AL806512
	16 719.803	31.6	135750	1	AL928859
C	17 719.803	31.6	142529	1	AC115852
	18 715.001	31.4	153772	1	AC129304
	19 715.001	31.4	271609	1	AC122210
C	20 714.803	31.4	247119	1	AC126868
	21 714.803	31.4	259629	1	AC119553
	22 713.403	31.3	235196	1	AC132786
	23 710.203	31.2	247726	1	AC128004
C	24 710.102	31.2	245174	1	AC095157
	25 707.202	31.0	167223	1	AC113514
	26 706.902	31.0	250188	1	AC110100
	27 706.103	31.0	168169	1	AC129113
	28 704.103	30.9	198631	1	AL683804
C	29 704.103	30.9	267360	1	AC101556
	30 703.903	30.9	128936	1	AC015893
C	31 702.902	30.8	211245	1	AC126499
	32 702.902	30.8	241255	1	AC095214
	33 699.902	30.7	227762	1	AL669976

RESULT 1

AC124438
LOCUS
DEFINITION
Mus musculus BAC clone RP24-260K9 from chromosome 7, complete sequence.
AC124438
AC124438.4 GI:23462943
HTG.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
McLellan, M., Bielicki, L., Dignan, G. and Doeber, A.
1 (bases 1 to 183326)
The sequence of Mus musculus BAC clone RP24-260K9
Unpublished (2001)
2 (bases 1 to 183326)
Wilson, R.
Sequencing of Mus musculus
Unpublished (2001)
3 (bases 1 to 183326)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 183326)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (01-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 183326)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (02-OCT-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 183326)
Wilson, R.
Direct Submission
Submitted (13-NOV-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 2, 2002 this sequence version replaced gi:22539265.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu

Summary Statistics
Center project name: M_BB0260K09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source

```
1. 183326
   Location/Qualifiers
     /organism="Mus musculus"
     /mol_type="genomic DNA"
     /db_xref="taxon:10090"
     /chromosome="7"
     /map="7"
     /clone="RP24-260K9"
     /clone_lib="RPCI-24"
```

```
421..527
   rpt_family="L1"
   repeat_region
780..849
   rpt_family="L1"
   repeat_region
1229..1489
   rpt_family="L1"
   repeat_region
1925..2147
   rpt_family="MER1_type"
   repeat_region
2269..2345
   rpt_family="B4"
   repeat_region
2432..2639
   rpt_family="L1"
   repeat_region
2657..2747
   rpt_family="L1"
   repeat_region
2753..2819
   rpt_family="ID"
   repeat_region
2825..2891
   rpt_family="L1"
   repeat_region
3225..4557
   rpt_family="L1"
   repeat_region
5567..5680
   rpt_family="MIR"
   repeat_region
6490..6682
   rpt_family="B2"
   repeat_region
6777..6881
   rpt_family="Alu"
   repeat_region
6889..7045
   rpt_family="B4"
   repeat_region
9156..9335
   rpt_family="MER2_type"
   repeat_region
10113..10156
   rpt_family="B4"
   repeat_region
10440..10701
   rpt_family="B4"
   repeat_region
11045..11114
   rpt_family="ID"
   repeat_region
11205..11581
   rpt_family="L1"
   repeat_region
12639..12861
   rpt_family="B4"
   repeat_region
12916..13988
   rpt_family="B4"
   repeat_region
13314..13598
   rpt_family="MaLR"
```

```
repeat_region
14985..15253
   rpt_family="B4"
   repeat_region
15284..15355
   rpt_family="MIR"
   repeat_region
15669..15726
   rpt_family="ERV1"
   repeat_region
16139..16229
   rpt_family="L1"
   repeat_region
16418..16525
   rpt_family="L1"
   repeat_region
16880..16932
   rpt_family="MIR"
   repeat_region
16942..17139
   rpt_family="L1"
   repeat_region
17302..17445
   rpt_family="Alu"
   repeat_region
17459..17508
   rpt_family="L1"
   repeat_region
17667..17787
   rpt_family="Alu"
   repeat_region
18127..18260
   rpt_family="L2"
   repeat_region
18386..18472
   rpt_family="MIR"
   repeat_region
18473..18818
   rpt_family="MaLR"
   repeat_region
18831..19229
   rpt_family="MaLR"
   repeat_region
19435..19530
   rpt_family="MIR"
   repeat_region
20110..20262
   rpt_family="B4"
   repeat_region
20395..20447
   rpt_family="ERV1"
   repeat_region
20458..20581
   rpt_family="MaLR"
   repeat_region
21632..21854
   rpt_family="MER1_type"
   repeat_region
21927..22288
   rpt_family="MaLR"
   repeat_region
23000..23102
   rpt_family="MIR"
   repeat_region
23289..23480
   rpt_family="B2"
   repeat_region
23481..23670
   rpt_family="B2"
   repeat_region
23989..24034
   rpt_family="L2"
   repeat_region
24461..24730
   rpt_family="B4"
   repeat_region
25760..25970
   rpt_family="MER1_type"
   repeat_region
26408..26524
   rpt_family="Alu"
   repeat_region
26748..27090
   rpt_family="MaLR"
   repeat_region
28283..28734
   rpt_family="ERV1"
   repeat_region
29104..29319
   rpt_family="MER2_type"
   repeat_region
29316..29425
   rpt_family="MER2_type"
   repeat_region
29597..29747
   rpt_family="MIR"
   repeat_region
29825..29839
   rpt_family="MIR"
   repeat_region
30206..30399
   rpt_family="MER1_type"
   repeat_region
31148..31285
   rpt_family="MER1_type"
   repeat_region
32032..32336
   rpt_family="MaLR"
```


Query Match		36.9%;	Score	840.902;	DB	1;	Length	183326;
Best Local Similarity		68.6%;	Pred.	No.	0.024;			
Matches		1486;	Conservative	0;	Mismatches	456;	Indels	223;
Gaps		26;						
Qy	294	TATTTCTGATTTCTATCTTTGGCTCAATTTTAACTCAGTAGTGAGTGTGTTGGTTCCAT	353					
Db	40783	TCCTTCTAAATTTCTACCTTGACTCAATTTTCAATCTATAATGAGTATTTTAGTTTCCAT	40842					
Qy	354	AAGTTTGTAAGTTTCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	413					
Db	40843	GAATTTGTAACTTTC-----TTGTTGTTGATATCCAGCTTTTAAGCAATGG	40887					
Qy	414	TGTFACAGATAGGACATAGATATTAATTCAAATTTGCTTTTATCTGFCGAGACTTGCTTTG	473					
Db	40888	TGGGATAATAGTGTGCGAGGAGTATTTTCAATTTTCTTGTATCTGTGAGATTTGCTTTG	40947					
Qy	474	TTTTGAAATATGATATCAATTTTGGAGAGTTTTCATAGGGTGTGACAAAGAGTACAGTC	533					
Db	40948	TGCTGAGTATGTGTCAATTTTGGAGAGTTTC--CATGAGCTGCTGAGAGGATATATTT	41005					
Qy	534	TTTCTGTTTTGGTGAATAGTCTGTAAATATCTAGTCCACTTGGTTTATGACATCAG	593					
Db	41006	TTT-TTGTGTTCAAGTGAAATGTTCTACAAATATTTTAGGTCCATTTGTTTCTAAATGAAG	41065					
Qy	594	TTAGCTCCAGCATTTCTCTGTTTCTGTTTGTGTTGTTGAGATGACCTACTGTTGGAGAGAT	653					
Db	41066	TTAGCTCCAGTAGATACCTGTTTAGCTTTTGTCTGA--ACAACCTATTTATCGGCAAGAAT	41124					
Qy	654	GGGGTATGAGTAGGCCACTATCTGTGTGTGAGG--TCAATATGTCAATTTAGCTGTAGC	712					
Db	41125	GGAGTATGAAATCACCACCTAACACTGTGTGAGGATCAATATGTGACITTAGCTGTAGT	41184					
Qy	713	TGTCCTTGTTTTATGAACCTGGGTGACATTTGTTGTTGGTGCATAGACATTAAGAAATGCA	772					
Db	41185	AGTGTCTTTTATGAACCTTGGATGACCTTGTTTTCGGTGCAGAGATAATAAGAAAGGCA	41244					
Qy	773	ATGTCTCTTGGTGA--TTTTCCTTTGATGCTATGTAGTATTTCTTCCCAATCTCATC	829					
Db	41245	ATATCTCTTTATATAATTTTTCCTTTTGATATTTATGTAGTGTCTTTCCTATCTCTTC	41304					
Qy	830	TGCTTAGTTTGGGTTTAAGTCTAT--TAGTCAGATATATAAATGACTGTAT--CGGCTTG	886					
Db	41305	TGAATAGTTTGTATTTGAAGTCTGTCTCATCAAAATAGTAAATGACTATATACAGCTTG	41364					
Qy	887	CTTCTTAGGGCCATTTGCTTAGATATC--TTTTCCTCTCTTTTACTCTAGGTGATGTCT	945					
Db	41365	CTTGTAGGTCATTTGCTTCAAAATATCTTTTCCATCTTTTACTCTTAGGTAATGACT	41424					
Qy	946	ATCC--ATGTAGTGTGTCTTTTGGATGACAGCAGTAGGATGGAT--CTTGTTTTCATA	1001					
Db	41425	ATCCTTGATGTCAGGTGCAATTTCTTGAATACACAGCAAGAATGGATCCCTGTTTTCACA	41484					
Qy	1002	TCCATCTGTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCATGTATGTT	1061					
Db	41485	CTTATCTGTGTAGGCTATGTGTTTTTACTGGGTAACTAAGA-----CTACTGATCTT	41536					
Qy	1062	GAGAAATCAATGAGCAGTGTGTTGGATTCCTTTTATCTTGCACTTGTGAAGTGTGTG	1121					
Db	41537	AAAAAGCTTAATGAGCAGTCTCTGTGATCTCTGCTCTTTTCTCTGTGTGTGTG	41596					
Qy	1122	TGTFGTG	1181					
Db	41597	TCTG	41656					
Qy	1182	TGTCGTGTCT-----	1192					
Db	41657	TGTGTCTGTCTGCT	41716					
Qy	1193	-----	1192					
Db	41717	TGTCGTCTCTGCT	41776					


```

Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
McElrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramaasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (31-JUL-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 192016)
Birren,B., Nusbaum,C. and Lander,E.
Direct Submission
Submitted (19-AUG-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 192016)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukgafter,B., Camarata,J., Chang,J., Choepel,Y.,
Collamore,A., Cook,A., Cooke,P., Corum,B., DeAtellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,B., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Liu,X., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramaasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-SEP-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 19, 2003 this sequence version replaced gi:33354396.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: Li8775
Center clone name: 88_A_10
-----
Location/Qualifiers
1. .192016
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="1"
/map="1"
/clone="RP24-88A10"
/clone_lib="RP24-88A10"
1. .28181
/note="wgs end extension"
clone end:SP6"
3. .3738
/rpt_family="RLTRETN_MM-int"
complement(1432..1436)
/note="<30 qual SINGL region"
complement(1454..1817)

misc_feature
1. .28181
/note="wgs end extension"
clone end:SP6"
3. .3738
/rpt_family="RLTRETN_MM-int"
complement(1432..1436)
/note="<30 qual SINGL region"
complement(1454..1817)

repeat_region
1. .28181
/note="wgs end extension"
clone end:SP6"
3. .3738
/rpt_family="RLTRETN_MM-int"
complement(1432..1436)
/note="<30 qual SINGL region"
complement(1454..1817)

unsure
1. .28181
/note="wgs end extension"
clone end:SP6"
3. .3738
/rpt_family="RLTRETN_MM-int"
complement(1432..1436)
/note="<30 qual SINGL region"
complement(1454..1817)

unsure
1. .28181
/note="wgs end extension"
clone end:SP6"
3. .3738
/rpt_family="RLTRETN_MM-int"
complement(1432..1436)
/note="<30 qual SINGL region"
complement(1454..1817)

```


QY 2230 TTGGCATATAGA 2241
DB 13242 TTTTGAAAGA 13253

RESULT 5
AC102686/c
LOCUS AC102686.2 GI:30018186
DEFINITION HTG: HTGS PHASE1; HTGS DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Mus musculus, clone RP24-467J24
JOURNAL Unpublished
AUTHORS

REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckghalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckghalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataranan, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Apr 17, 2003 this sequence version replaced gi:17060796.
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L19265
Center clone name: 467_J_24

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 1571: contig of 1571 bp in length
* 1572 1671: gap of 100 bp
* 1672 7052: contig of 5381 bp in length
* 7053 7152: gap of 100 bp
* 7153 93807: contig of 8655 bp in length
* 93808 93907: gap of 100 bp
* 93908 227550: contig of 133643 bp in length.
Location/Qualifiers
1. 227550
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-467J24"
/clone_lib="RP24-467J24"
1. 1571
/note="assembly_fragment"
1672. 7052
/note="assembly_fragment"
7153. 93807
/note="assembly_fragment"
93908. 227550
/note="assembly_fragment"

Query Match 34.1%; Score 778.202; DB 1; Length 227550;
Best Local Similarity 69.7%; Pred. No. 0.062;
Matches 1430; Conservative 0; Mismatches 508; Indels 114; Gaps 29;

QY 277 TCTCTTCCAAACACTTCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTG 336
DB 157107 TTTTCATTCATTTCTTGAAGTCTTTAAATTTCTTCTTGACACATTTTCATTCATTAGAC 157048

QY 337 AGTTGTTGGTTTCCATAAAGTTTGTAGTTTCTGTGTTTCTGTGTTTCTGTGTTTCTGTGTTTAT 396
DB 157047 AGTTGTTTGTAGTTTCCATGAGTTTGTAGGCTTTCTCCTGATTC---TGTTTGTTCATAT 156991

QY 397 CTAGATTAGCTGTGCTGTGATAGAGGA--CATAGAGTATTATTTCATTTGCTCTTTTA 454
DB 156990 CTGGCTTTTATACATGGTAGTACAGATGAGATGAGCATATTTTGGTTTCTTTGTA 156931

QY 455 TCTGTCGAGACTTGC-TTTGTTTGTAAATATGATTTCAATTTG--GAGAGTTTCATAGG 511
DB 156930 TCTGTTGAACTTCTTTTGTGATTGAAGATGCTTCAATTTTGAAGAAAGTTTCATGAG 156871

QY 512 GTCTGACAGAGAGTACAGTC-TTGTCTGTTTGTGTAATAGTCTGTAATATCT-CTA 569
DB 156870 GTCTGAGAAAAGGATCTCTTTTGTGTTGGGTAATATGTTCTGTAATATGTTGA 156811

QY 570 GGTCCACTTGGTTTATGACATCAGTTAGTCCAGCATTTCTCTGTTTCGTTTTTTGTTGA 629
DB 156810 GGTACAGTTCAATTTATAACATAGTTAGTCTAGCTCTGC-----TTAATTTTGTCTG 156757

QY 630 GATGACCTTAACCTTTGGAGAGAAATGGGGTATTGAAGTACCCACTATCTGTGTGTCAG-G 688
DB 156756 GATGACCTTCCTAGAGGCAAGAGTTGGGAATTAAGTCTCCCACTCTCAGTGCATGAGTG 156697

Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 217398)

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collumore, A., Cook, A., Cooke, P., Corum, B., DeArelano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Harez, N., Hagopian, D., Hegos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Melarim, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
Submitted (09-JUN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 9, 2003 this sequence version replaced gi:22380823.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RN/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 560.19
Center clone name: L18952

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 206419 bases at least Q40
Consensus quality: 21145 bases at least Q30
Consensus quality: 212661 bases at least Q20
Insert size: 182000; agarose-fp
Insert size: 213698; sum-of-contigs
Quality coverage: 12.1 in Q20 bases; agarose-fp
Quality coverage: 10.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 38 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 19129: contig of 19129 bp in length
* 19130 19229: gap of 100 bp
* 19230 19836: contig of 607 bp in length
* 19837 19936: gap of 100 bp
* 19937 20613: contig of 677 bp in length
* 20614 20713: gap of 100 bp
* 20714 21355: contig of 642 bp in length

FEATURES
source

Location/Qualifiers
1. .217398
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

21356 gap of 100 bp
21455: contig of 975 bp in length
22430: gap of 100 bp
22530: contig of 615 bp in length
23146: contig of 100 bp
23246: gap of 100 bp
23990: contig of 744 bp in length
24090: gap of 100 bp
24091: contig of 567 bp in length
24557: gap of 100 bp
24757: contig of 720 bp in length
25577: gap of 100 bp
26032: contig of 515 bp in length
26092: gap of 100 bp
26192: contig of 644 bp in length
26936: gap of 100 bp
27400: contig of 464 bp in length
27500: gap of 100 bp
28196: contig of 696 bp in length
28296: gap of 100 bp
28967: contig of 671 bp in length
29067: gap of 100 bp
29828: contig of 761 bp in length
29928: gap of 100 bp
30975: contig of 1047 bp in length
31075: gap of 100 bp
31970: contig of 895 bp in length
32070: gap of 100 bp
33032: contig of 962 bp in length
33132: gap of 100 bp
33910: contig of 778 bp in length
34010: gap of 100 bp
34616: contig of 606 bp in length
34716: gap of 100 bp
35694: contig of 978 bp in length
35794: gap of 100 bp
36680: contig of 886 bp in length
36780: gap of 100 bp
37111: contig of 931 bp in length
37811: gap of 100 bp
38862: contig of 1051 bp in length
38962: gap of 100 bp
39805: contig of 843 bp in length
39905: gap of 100 bp
41567: contig of 1662 bp in length
41667: gap of 100 bp
43156: contig of 1489 bp in length
43256: gap of 100 bp
44564: contig of 1308 bp in length
44664: gap of 100 bp
46143: contig of 1479 bp in length
46243: gap of 100 bp
47645: contig of 1402 bp in length
47745: gap of 100 bp
49089: contig of 1324 bp in length
49169: gap of 100 bp
51131: contig of 1962 bp in length
51231: gap of 100 bp
52967: contig of 1736 bp in length
53067: gap of 100 bp
56770: contig of 3703 bp in length
56870: gap of 100 bp
81259: contig of 24389 bp in length
81359: gap of 100 bp
114087: contig of 32728 bp in length
114187: gap of 100 bp
158558: contig of 44371 bp in length
158559 gap of 100 bp
158659 contig of 58740 bp in length.

[illegible]

```

RESULT 7
AL824706
LOCUS
DEFINITION
AL824706 207687 bp DNA linear ROD 17-APR-2003
Mouse DNA sequence from clone RP23-218A4 on chromosome 4, complete
sequence.
ACCESSION
AL824706
VERSION
AL824706.8 GI:30024422
KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 207687)
AUTHORS
Tracey, A.
TITLE
Direct Submission
JOURNAL
Submitted (17-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerv@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Apr 16, 2003 this sequence version replaced gi:25045345.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Contact: humquerv@sanger.ac.uk
Web site: http://www.sanger.ac.uk
Center code: SC

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> RP23-218A4 is from the RPCI-23 Mouse BAC Library constructed by the group of Piter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3.6.

FEATURES

Location/Qualifiers

source	1. .207687	
	/organism="Mus musculus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10090"	
	/chromosome="4"	
	/clone="RP23-218A4"	
	/clone_lib="RPC1-23"	
Query Match	33.4%; Score 761.303; DB 1; Length 207687;	
Best Local Similarity	68.6%; Pred. No. 0.071;	
Matches 1387; Conservative	0; Mismatches 477; Indels 159; Gaps 26;	
QY	300 TTGATTTCTATCTTTGGCTCA---TTTTAACTCAGTAGTGAGTTGTTGGTTTCCATAAG 356	
Db	114106 TTCAATTCCTTTCTTGACCCATTTTTTTCATTCAGTAGTGAATTGTTCCATGCTGTGAAG 114165	
QY	357 TTTGAAAGTTTTCGTGTTCTCTGTTGTTGTTGTTTATCTAGATTAAAGCTGTGGTGG 416	
Db	114166 TTTGTAAACTTTCTGAGGTTTCTGCTGGTGGTATAT--CCCACTTTAAATCCACGTGG 114223	
QY	417 TCAGATGAGCATAGAGTATTATTCAAATGCTTTTATCTGTCGAGACTGCTGTTGTTT 476	
Db	114224 TCAGATGAGTCAGGCGGTATTATTGATTTCTTCTATCTATTGAGACTTCTCTTTGTT 114283	
QY	477 TGAATATGTATTCAATTTTTGGAGAGTTTCATAGGTGCTGCAAGAAGGTACAGTCTTT 536	
Db	114284 CTGAGCATATGGTGTATTTTGGAGAAAGTTCTCAGGTGCTGAGAAAAAG-----TGT 114336	
QY	537 GTGTTTGGTGAATAGTCTGTAATAATCTCTAGGTCCACTTGGTTTATGATCAGATT 596	
Db	114337 ATGTTTGTTGAAAAAGGTTCTGTTAATA-GGCTAGGTCTTTATGGTTTATAGCCTCAGTT 114395	
QY	597 GCTCCAGCAATTTCTCTGTTTCTGTTTGTGTAGATGACCTAACTGTTGGAGAGAAATGGG 656	
Db	114396 GCTCCAGCAATTTCTATTTAGTTTGTCTG-GATGACCTGCTATTGGTGAGAGTGA 114454	
QY	657 GTATTGAAGTAGCCCATCTCTGTGTGTGAGG-TCATAATGATTTTACGTGTAGCTGT 715	
Db	114455 GTATTGAAGTCTCCCATGTCAGCAGTGTGTGAGGCTCAATATGTGAATTAAGCCCTAGCAGT 114514	
QY	716 GCTGTTTATGAACCTTGGGT-GACATTTGTTTGGTGTGATACACAAATAAGAATTGCAAT 774	
Db	114515 ATTTCTCTATGAACCTTGGGTAACCTTGTGTTTGGTGATAGCTCTTGAACTTGCAAT 114574	
QY	775 GTCCTCTTTGGTGA-TTTTCTTTGATGCCATGTATGTATCTTTCCCAATCTCATCTGT 833	
Db	114575 GTCCTCTTTGGTGTACTTTTCTTTGATGACTATGCAAGTGCCTTCCCTATCTCTTGAT 114634	
QY	834 TAGTTTGGGTTTAAGTCTATTAG-TCAGATATTAAATGACATGATCCGCTTGCTTCTT 892	
Db	114635 TAGTTTGGGTTGAAGTGTGTATTTTAAAAATTAATAAGGCTACACAGCTTGTCTTCTT 114694	
QY	893 AGGGCCATTTGCTTAGAATATCTTTCCATCCTTTTACTCTAAGGTGATGTCTATCCATG 952	
Db	114695 AGGTCCATTTGCTTGTAAATACCAATTT--TCCATCCACCCTGGGTAATGTCTACCCCTG 114751	
QY	953 ---GTAGTGTCTTTTTTGGATCCAGCAGTAGGATGGATCTTGTTTTTCATATCCATCTCT 1009	
Db	114752 CTATTAAATGCTGATTTCTTGGATCAACAGAAAGATGGATTCGCGTTTACATCCATCTT 114811	
QY	1010 GTTACCAGTATCTTTTCTAGAGAAATTAAGATCAATTGAGTCAATTCATGTTGAGAAATTA 1069	
Db	114812 ATTAGTCTGTCTTTTATATGGGAATTTGA-----GACCACAGATGTTGAGAGATA 114862	
QY	1070 TCAATGAGCAGGTTTGTGGAAATCTGTGTTATCTTGCACCTGTGAAGTGTGTGTGTGTG 1129	
Db	114863 TCAATGAATGATGATGTAATTCCTGTAT-----TGTAAGTGTGTATGTGT--- 114909	
QY	1130 TG 1189	
Db	114910 -----TGATATGTGTGTGTATTTG 114929	
QY	1190 TCTTGTGTGTGTGTGTCTCTCCCTCTTTTGTATTTTGGCTGGGAATTAATTAATTAATTC 1249	

Talimas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 17, 2003 this sequence version replaced gi:28604153.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L25752
 Center clone name: 492_C_11

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 18 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 50033: contig of 50033 bp in length
 * 50034 50133: gap of 100 bp
 * 50134 50435: contig of 302 bp in length
 * 50436 50536: gap of 100 bp
 * 50536 51197: contig of 662 bp in length
 * 51198 51297: gap of 100 bp
 * 51298 52656: contig of 1359 bp in length
 * 52657 52756: gap of 100 bp
 * 52757 57009: contig of 4253 bp in length
 * 57010 57109: gap of 100 bp
 * 57110 60822: contig of 3713 bp in length
 * 60823 60923: gap of 100 bp
 * 60923 66293: contig of 5371 bp in length
 * 66294 66393: gap of 100 bp
 * 66394 72812: contig of 6419 bp in length
 * 72813 72913: gap of 100 bp
 * 72913 82696: contig of 9784 bp in length
 * 82697 94299: contig of 11503 bp in length
 * 94300 94399: gap of 100 bp
 * 94400 118094: contig of 23695 bp in length
 * 118095 143491: gap of 100 bp
 * 143492 143591: contig of 25297 bp in length
 * 143592 167313: gap of 100 bp
 * 167313 167412: contig of 23721 bp in length
 * 167413 180073: contig of 12661 bp in length
 * 180074 180174: gap of 100 bp
 * 180174 198870: contig of 18597 bp in length
 * 198871 261067: gap of 100 bp
 * 261067 302119: contig of 62096 bp in length
 * 302119 302120: contig of 40953 bp in length
 * 302120 343010: gap of 100 bp
 * 302220 343010: contig of 40791 bp in length.

FEATURES

source
 1..343010
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP24-492C11"
 /clone_lib="RPCI-24 Male Mouse BAC"
 1..50033
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 misc_feature
 1..50033

misc_feature 50134..50435
 /note="assembly_fragment"
 misc_feature 50536..51197
 /note="assembly_fragment"
 misc_feature 51298..52656
 /note="assembly_fragment"
 misc_feature 52757..57009
 /note="assembly_fragment"
 misc_feature 57110..60822
 /note="assembly_fragment"
 misc_feature 60923..66293
 /note="assembly_fragment"
 misc_feature 66394..72812
 /note="assembly_fragment"
 misc_feature 72913..82696
 /note="assembly_fragment"
 misc_feature 82797..94299
 /note="assembly_fragment"
 misc_feature 94400..118094
 /note="assembly_fragment"
 misc_feature 118195..143491
 /note="assembly_fragment"
 misc_feature 143592..167312
 /note="assembly_fragment"
 misc_feature 167413..180073
 /note="assembly_fragment"
 misc_feature 180174..198870
 /note="assembly_fragment"
 misc_feature 198971..261066
 /note="assembly_fragment"
 misc_feature 261167..302119
 /note="assembly_fragment"
 misc_feature 302220..343010
 /note="assembly_fragment"
 Query Match 33.0%; Score 752.402; DB 1; Length 343010;
 Best Local Similarity 70.5%; Pred. No. 0.12;
 Matches 1426; Conservative 0; Mismatches 456; Indels 140; Gaps 33;
 Qy 300 TTGATTTCTATCTTGGCTCA--TTTAACTCAGTAGTGTGTTGGTTTCCATAAGT 357
 Db 75922 TTGATTTCTTCTTGACCAATTTTTCATTCAGTAGTGTGTTTCCATTTCCATGAGT 75863
 Qy 358 TTGTAAGTTTCTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 417
 Db 75862 TTGTAAGTTTCTGTTGTT--TTATTTGTTGATCTCAGCTTTGATCTGTGGTGT 75807
 Qy 418 CAGATAGGACATAGAGTATTATTTCAATTGTTCTTTTATCTGTGAGACTTGTCTTTT 477
 Db 75806 CAGATAGTATGCGGGGACTATTTCATTTTCTTCTGTGTTGAAACCTGCTCTGATC 75747
 Qy 478 GAAATATGATTCAATTTTGGAG--AGTTTCTAGGTGCTGACAAAGAGGTACAGTCTT 535
 Db 75746 TGAATATAGCTCAGTTTGGAGAAAGTTCCATAGTACTGAGAGAAAGTATTCTT 75687
 Qy 536 TGTGTTTGTGAAATAGTCTGTAATATCT-CTAGGTCCACTTGGTTTATGACATCAGT 594
 Db 75686 TGTGTTGGGTGAAATATTCTGTAATATTATTAGGTCTATTGTTGTTTATAATGTTT 75627
 Qy 595 TAGCTCCAGCATTTCTGTTTCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 654
 Db 75626 TAACCTCAGTATTCTC-----TTTAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 75574
 Qy 655 GGGTATTGAAAGTAGCCCACTATCTGTGTGTGAGGTCAATATGTTAGTTAGTGTGTGCTG 714
 Db 75573 GAGTATTGAAAGTCTCCCACTGTGAGAATGTAAG-----TATCAATTAGTAC 75528
 Qy 715 TGCTTGTGTTTATGAACCTTGGGTGACATTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 774
 Db 75527 TGTTCCTTTTTCGAACCTTAGGTTCCTCCCTTGTGTTGTTGTTGTTGTTGTTGTTGTT 75469
 Qy 775 GTCTCTTGTGTGGA--TTTTCCTTTGTATGCTCTATGTAGTATTCTTCCCAATCTCATCTGT 833

AUTHORS Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C., MacDonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vasiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wymann,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (24-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Jan 24, 2003 this sequence version replaced gi:15947185. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information -----
Center project name: L17321
Center clone name: 465_J20

----- Summary Statistics -----
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 228037 bases at least Q40
Consensus quality: 231924 bases at least Q30
Consensus quality: 233041 bases at least Q20
Insert size: 210000; agarose-
Insert size: 234204; sum-of-
Quality coverage: 11.8 in Q20 bases; agarose-
Quality coverage: 10.6 in Q20 bases; sum-of-
contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 84: contig of 84 bp in length
85 184: gap of 100 bp
185 980: contig of 796 bp in length
981 1080: gap of 100 bp
1081 1727: contig of 647 bp in length
1728 1827: gap of 100 bp
1828 2827: contig of 700 bp in length
2828 3617: contig of 990 bp in length
3618 3717: gap of 100 bp
3718 4707: contig of 990 bp in length
4708 4807: gap of 100 bp
4808 5740: contig of 933 bp in length
5741 6338: contig of 698 bp in length
6339 7303: contig of 665 bp in length
7304 7403: gap of 100 bp

* 7404 8266: contig of 863 bp in length
8267 8366: gap of 100 bp
8367 9570: contig of 1284 bp in length
9571 11332: contig of 1562 bp in length
11332 12137: contig of 805 bp in length
12137 12237: gap of 100 bp
12237 13117: contig of 880 bp in length
13117 13884: contig of 667 bp in length
13884 13984: gap of 100 bp
13984 15206: contig of 1222 bp in length
15206 15306: gap of 100 bp
15306 16631: contig of 1325 bp in length
16631 16731: gap of 100 bp
16731 18149: contig of 1418 bp in length
18149 18249: gap of 100 bp
18249 19571: contig of 1322 bp in length
19571 19671: gap of 100 bp
19671 27589: contig of 7918 bp in length
27589 27689: gap of 100 bp
27689 32790: contig of 111301 bp in length
32790 33991: gap of 100 bp
33991 364589: contig of 25499 bp in length
364589 364689: gap of 100 bp
364689 364690: contig of 19366 bp in length
364690 364691: gap of 100 bp
364691 364692: gap of 100 bp
364692 364693: gap of 100 bp
364693 364694: gap of 100 bp
364694 364695: gap of 100 bp
364695 364696: gap of 100 bp
364696 364697: gap of 100 bp
364697 364698: gap of 100 bp
364698 364699: gap of 100 bp
364699 364700: gap of 100 bp
364700 364701: gap of 100 bp
364701 364702: gap of 100 bp
364702 364703: gap of 100 bp
364703 364704: gap of 100 bp
364704 364705: gap of 100 bp
364705 364706: gap of 100 bp
364706 364707: gap of 100 bp
364707 364708: gap of 100 bp
364708 364709: gap of 100 bp
364709 364710: gap of 100 bp
364710 364711: gap of 100 bp
364711 364712: gap of 100 bp
364712 364713: gap of 100 bp
364713 364714: gap of 100 bp
364714 364715: gap of 100 bp
364715 364716: gap of 100 bp
364716 364717: gap of 100 bp
364717 364718: gap of 100 bp
364718 364719: gap of 100 bp
364719 364720: gap of 100 bp
364720 364721: gap of 100 bp
364721 364722: gap of 100 bp
364722 364723: gap of 100 bp
364723 364724: gap of 100 bp
364724 364725: gap of 100 bp
364725 364726: gap of 100 bp
364726 364727: gap of 100 bp
364727 364728: gap of 100 bp
364728 364729: gap of 100 bp
364729 364730: gap of 100 bp
364730 364731: gap of 100 bp
364731 364732: gap of 100 bp
364732 364733: gap of 100 bp
364733 364734: gap of 100 bp
364734 364735: gap of 100 bp
364735 364736: gap of 100 bp
364736 364737: gap of 100 bp
364737 364738: gap of 100 bp
364738 364739: gap of 100 bp
364739 364740: gap of 100 bp
364740 364741: gap of 100 bp
364741 364742: gap of 100 bp
364742 364743: gap of 100 bp
364743 364744: gap of 100 bp
364744 364745: gap of 100 bp
364745 364746: gap of 100 bp
364746 364747: gap of 100 bp
364747 364748: gap of 100 bp
364748 364749: gap of 100 bp
364749 364750: gap of 100 bp
364750 364751: gap of 100 bp
364751 364752: gap of 100 bp
364752 364753: gap of 100 bp
364753 364754: gap of 100 bp
364754 364755: gap of 100 bp
364755 364756: gap of 100 bp
364756 364757: gap of 100 bp
364757 364758: gap of 100 bp
364758 364759: gap of 100 bp
364759 364760: gap of 100 bp
364760 364761: gap of 100 bp
364761 364762: gap of 100 bp
364762 364763: gap of 100 bp
364763 364764: gap of 100 bp
364764 364765: gap of 100 bp
364765 364766: gap of 100 bp
364766 364767: gap of 100 bp
364767 364768: gap of 100 bp
364768 364769: gap of 100 bp
364769 364770: gap of 100 bp
364770 364771: gap of 100 bp
364771 364772: gap of 100 bp
364772 364773: gap of 100 bp
364773 364774: gap of 100 bp
364774 364775: gap of 100 bp
364775 364776: gap of 100 bp
364776 364777: gap of 100 bp
364777 364778: gap of 100 bp
364778 364779: gap of 100 bp
364779 364780: gap of 100 bp
364780 364781: gap of 100 bp
364781 364782: gap of 100 bp
364782 364783: gap of 100 bp
364783 364784: gap of 100 bp
364784 364785: gap of 100 bp
364785 364786: gap of 100 bp
364786 364787: gap of 100 bp
364787 364788: gap of 100 bp
364788 364789: gap of 100 bp
364789 364790: gap of 100 bp
364790 364791: gap of 100 bp
364791 364792: gap of 100 bp
364792 364793: gap of 100 bp
364793 364794: gap of 100 bp
364794 364795: gap of 100 bp
364795 364796: gap of 100 bp
364796 364797: gap of 100 bp
364797 364798: gap of 100 bp
364798 364799: gap of 100 bp
364799 364800: gap of 100 bp
364800 364801: gap of 100 bp
364801 364802: gap of 100 bp
364802 364803: gap of 100 bp
364803 364804: gap of 100 bp
364804 364805: gap of 100 bp
364805 364806: gap of 100 bp
364806 364807: gap of 100 bp
364807 364808: gap of 100 bp
364808 364809: gap of 100 bp
364809 364810: gap of 100 bp
364810 364811: gap of 100 bp
364811 364812: gap of 100 bp
364812 364813: gap of 100 bp
364813 364814: gap of 100 bp
364814 364815: gap of 100 bp
364815 364816: gap of 100 bp
364816 364817: gap of 100 bp
364817 364818: gap of 100 bp
364818 364819: gap of 100 bp
364819 364820: gap of 100 bp
364820 364821: gap of 100 bp
364821 364822: gap of 100 bp
364822 364823: gap of 100 bp
364823 364824: gap of 100 bp
364824 364825: gap of 100 bp
364825 364826: gap of 100 bp
364826 364827: gap of 100 bp
364827 364828: gap of 100 bp
364828 364829: gap of 100 bp
364829 364830: gap of 100 bp
364830 364831: gap of 100 bp
364831 364832: gap of 100 bp
364832 364833: gap of 100 bp
364833 364834: gap of 100 bp
364834 364835: gap of 100 bp
364835 364836: gap of 100 bp
364836 364837: gap of 100 bp
364837 364838: gap of 100 bp
364838 364839: gap of 100 bp
364839 364840: gap of 100 bp
364840 364841: gap of 100 bp
364841 364842: gap of 100 bp
364842 364843: gap of 100 bp
364843 364844: gap of 100 bp
364844 364845: gap of 100 bp
364845 364846: gap of 100 bp
364846 364847: gap of 100 bp
364847 364848: gap of 100 bp
364848 364849: gap of 100 bp
364849 364850: gap of 100 bp
364850 364851: gap of 100 bp
364851 364852: gap of 100 bp
364852 364853: gap of 100 bp
364853 364854: gap of 100 bp
364854 364855: gap of 100 bp
364855 364856: gap of 100 bp
364856 364857: gap of 100 bp
364857 364858: gap of 100 bp
364858 364859: gap of 100 bp
364859 364860: gap of 100 bp
364860 364861: gap of 100 bp
364861 364862: gap of 100 bp
364862 364863: gap of 100 bp
364863 364864: gap of 100 bp
364864 364865: gap of 100 bp
364865 364866: gap of 100 bp
364866 364867: gap of 100 bp
364867 364868: gap of 100 bp
364868 364869: gap of 100 bp
364869 364870: gap of 100 bp
364870 364871: gap of 100 bp
364871 364872: gap of 100 bp
364872 364873: gap of 100 bp
364873 364874: gap of 100 bp
364874 364875: gap of 100 bp
364875 364876: gap of 100 bp
364876 364877: gap of 100 bp
364877 364878: gap of 100 bp
364878 364879: gap of 100 bp
364879 364880: gap of 100 bp
364880 364881: gap of 100 bp
364881 364882: gap of 100 bp
364882 364883: gap of 100 bp
364883 364884: gap of 100 bp
364884 364885: gap of 100 bp
364885 364886: gap of 100 bp
364886 364887: gap of 100 bp
364887 364888: gap of 100 bp
364888 364889: gap of 100 bp
364889 364890: gap of 100 bp
364890 364891: gap of 100 bp
364891 364892: gap of 100 bp
364892 364893: gap of 100 bp
364893 364894: gap of 100 bp
364894 364895: gap of 100 bp
364895 364896: gap of 100 bp
364896 364897: gap of 100 bp
364897 364898: gap of 100 bp
364898 364899: gap of 100 bp
364899 364900: gap of 100 bp
364900 364901: gap of 100 bp
364901 364902: gap of 100 bp
364902 364903: gap of 100 bp
364903 364904: gap of 100 bp
364904 364905: gap of 100 bp
364905 364906: gap of 100 bp
364906 364907: gap of 100 bp
364907 364908: gap of 100 bp
364908 364909: gap of 100 bp
364909 364910: gap of 100 bp
364910 364911: gap of 100 bp
364911 364912: gap of 100 bp
364912 364913: gap of 100 bp
364913 364914: gap of 100 bp
364914 364915: gap of 100 bp
364915 364916: gap of 100 bp
364916 364917: gap of 100 bp
364917 364918: gap of 100 bp
364918 364919: gap of 100 bp
364919 364920: gap of 100 bp
364920 364921: gap of 100 bp
364921 364922: gap of 100 bp
364922 364923: gap of 100 bp
364923 364924: gap of 100 bp
364924 364925: gap of 100 bp
364925 364926: gap of 100 bp
364926 364927: gap of 100 bp
364927 364928: gap of 100 bp
364928 364929: gap of 100 bp
364929 364930: gap of 100 bp
364930 364931: gap of 100 bp
364931 364932: gap of 100 bp
364932 364933: gap of 100 bp
364933 364934: gap of 100 bp
364934 364935: gap of 100 bp
364935 364936: gap of 100 bp
364936 364937: gap of 100 bp
364937 364938: gap of 100 bp
364938 364939: gap of 100 bp
364939 364940: gap of 100 bp
364940 364941: gap of 100 bp
364941 364942: gap of 100 bp
364942 364943: gap of 100 bp
364943 364944: gap of 100 bp
364944 364945: gap of 100 bp
364945 364946: gap of 100 bp
364946 364947: gap of 100 bp
364947 364948: gap of 100 bp
364948 364949: gap of 100 bp
364949 364950: gap of 100 bp
364950 364951: gap of 100 bp
364951 364952: gap of 100 bp
364952 364953: gap of 100 bp
364953 364954: gap of 100 bp
364954 364955: gap of 100 bp
364955 364956: gap of 100 bp
364956 364957: gap of 100 bp
364957 364958: gap of 100 bp
364958 364959: gap of 100 bp
364959 364960: gap of 100 bp
364960 364961: gap of 100 bp
364961 364962: gap of 100 bp
364962 364963: gap of 100 bp
364963 364964: gap of 100 bp
364964 364965: gap of 100 bp
364965 364966: gap of 100 bp
364966 364967: gap of 100 bp
364967 364968: gap of 100 bp
364968 364969: gap of 100 bp
364969 364970: gap of 100 bp
364970 364971: gap of 100 bp
364971 364972: gap of 100 bp
364972 364973: gap of 100 bp
364973 364974: gap of 100 bp
364974 364975: gap of 100 bp
364975 364976: gap of 100 bp
364976 364977: gap of 100 bp
364977 364978: gap of 100 bp
364978 364979: gap of 100 bp
364979 364980: gap of 100 bp
364980 364981: gap of 100 bp
364981 364982: gap of 100 bp
364982 364983: gap of 100 bp
364983 364984: gap of 100 bp
364984 364985: gap of 100 bp
364985 364986: gap of 100 bp
364986 364987: gap of 100 bp
364987 364988: gap of 100 bp
364988 364989: gap of 100 bp
364989 364990: gap of 100 bp
364990 364991: gap of 100 bp
364991 364992: gap of 100 bp
364992 364993: gap of 100 bp
364993 364994: gap of 100 bp
364994 364995: gap of 100 bp
364995 364996: gap of 100 bp
364996 364997: gap of 100 bp
364997 364998: gap of 100 bp
364998 364999: gap of 100 bp
364999 365000: gap of 100 bp
365000 365001: gap of 100 bp
365001 365002: gap of 100 bp
365002 365003: gap of 100 bp
365003 365004: gap of 100 bp
365004 365005: gap of 100 bp
365005 365006: gap of 100 bp
365006 365007: gap of 100 bp
365007 365008: gap of 100 bp
365008 365009: gap of 100 bp
365009 365010: gap of 100 bp
365010 365011: gap of 100 bp
365011 365012: gap of 100 bp
365012 365013: gap of 100 bp
365013 365014: gap of 100 bp
365014 365015: gap of 100 bp
365015 365016: gap of 100 bp
365016 365017: gap of 100 bp
365017 365018: gap of 100 bp
365018 365019: gap of 100 bp
365019 365020: gap of 100 bp
365020 365021: gap of 100 bp
365021 365022: gap of 100 bp
365022 365023: gap of 100 bp
365023 365024: gap of 100 bp
365024 365025: gap of 100 bp
365025 365026: gap of 100 bp
365026 365027: gap of 100 bp
365027 365028: gap of 100 bp
365028 365029: gap of 100 bp
365029 365030: gap of 100 bp
365030 365031: gap of 100 bp
365031 365032: gap of 100 bp
365032 365033: gap of 100 bp
365033 365034: gap of 100 bp
365034 365035: gap of 100 bp
365035 365036: gap of 100 bp
365036 365037: gap of 100 bp
365037 365038: gap of 100 bp
365038 365039: gap of 100 bp
365039 365040: gap of 100 bp
365040 365041: gap of 100 bp
365041 365042: gap of 100 bp
365042 365043: gap of 100 bp
365043 365044: gap of 100 bp
365044 365045: gap of 100 bp
365045 365046: gap of 100 bp
365046 365047: gap of 100 bp
365047 365048: gap of 100 bp
365048 365049: gap of 100 bp
365049 365050: gap of 100 bp
365050 365051: gap of 100 bp
365051 365052: gap of 100 bp
365052 365053: gap of 100 bp
365053 365054: gap of 100 bp
365054 365055: gap of 100 bp
365055 365056: gap of 100 bp
365056 365057: gap of 100 bp
365057 365058: gap of 100 bp
365058 365059: gap of 100 bp
365059 365060: gap of 100 bp
365060 365061: gap of 100 bp
365061 365062: gap of 100 bp
365062 365063: gap of 100 bp
365063 365064: gap of 100 bp
365064 365065: gap of 100 bp
365065 365066: gap of 100 bp
365066 365067: gap of 100 bp
365067 365068: gap of 100 bp
365068 365069: gap of 100 bp
365069 365070: gap of 100 bp
365070 365071: gap of 100 bp
365071 365072: gap of 100 bp
365072 365073: gap of 100 bp
365073 365074: gap of 100 bp
365074 365075: gap of 100 bp
365075 365076: gap of 100 bp
365076 365077: gap of 100 bp
365077 365078: gap of 100 bp
365078 365079: gap of 100 bp
365079 365080: gap of 100 bp
365080 365081: gap of 100 bp
365081 365082: gap of 100 bp
365082 365083: gap of 100 bp
365083 365084: gap of 100 bp
365084 365085: gap of 100 bp
365085 365086: gap of 100 bp
365086 365087: gap of 100 bp
365087 365088: gap of 100 bp
365088 365089: gap of 100 bp
365089 365090: gap of 100 bp
365090 365091: gap of 100 bp
365091 365092: gap of 100 bp
365092 365093: gap of 100 bp
365093 365094: gap of 100 bp
365094 365095: gap of 100 bp
365095 365096: gap of 100 bp
365096 365097: gap of 100 bp
365097 365098: gap of 100 bp
365098 365099: gap of 100 bp
365099 365100: gap of 100 bp
365100 365101: gap of 100 bp
365101 365102: gap of 100 bp
365102 365103: gap of 100 bp
365103 365104: gap of 100 bp
365104 365105: gap of 100 bp
365105 365106: gap of 100 bp
365106 365107: gap of 100 bp
365107 365108: gap of 100 bp
365108 365109: gap of 100 bp
365109 365110: gap of 100 bp
365110 365111: gap of 100 bp
365111 365112: gap of 100 bp
365112 365113: gap of 100 bp
365113 365114: gap of 100 bp
365114 365115: gap of 100 bp
365115 365116: gap of 100 bp
365116 365117: gap of 100 bp
365117 365118: gap of 100 bp
365118 365119: gap of 100 bp
365119 365120: gap of 100 bp
365120 365121: gap of 100 bp
365121 365122: gap of 100 bp
365122 365123: gap of 100 bp
365123 365124: gap of 100 bp
365124 365125: gap of 100 bp
365125 365126: gap of 100 bp
365126 365127: gap of 100 bp
365127 365128: gap of 100 bp
365128 365129: gap of 100 bp
365129 365130: gap of 100 bp
365130 365131: gap of 100 bp
365131 365132: gap of 100 bp
365132 365133: gap of 100 bp
365133 365134: gap of 100 bp
365134 365135: gap of 100 bp
365135 365136: gap of 100 bp
365136 365137: gap of 100 bp
365137 365138: gap of 100 bp
365138 365139: gap of 100 bp
36

```
/note="assembly_fragment"
18250..19571
/note="assembly_fragment"
19672..27589
/note="assembly_fragment"
27690..138990
/note="assembly_fragment"
139091..164589
/note="assembly_fragment"
164690..184055
/note="assembly_fragment"
184156..234616
/note="assembly_fragment"

Query Match      33.0%; Score 752.302; DB 1; Length 236604;
Best Local Similarity 67.2%; Pred. No. 0.088;
Matches 1365; Conservative 0; Mismatches 502; Indels 163; Gaps 23;

Qy 297 TTCTGATTTCTATCTCGCTCAATTTTAACTCAGTAGTAGTGGTTTGGTTTCCATAAG 356
Db 73621 TTTCCAGTTCTGCTTGACCCATTTTTCCTCAATAGAGATTATTAGTTTCCATGAC 73680

Qy 357 TTGTGAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTATCTAGATTAAAGCTGTGGTG 416
Db 73681 TTTGTGAAGCTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 73735

Qy 417 TCAGATAGGATAGAGTATTATTCAATCTCTTTTATCTCTGCGAGACTGCTTCTGTT 476
Db 73736 TCAGATAGGATAGAGGTTTAAATTAAATTTCTGTTGTTGTTGTTGTTGTTGTTGTT 73795

Qy 477 TGAATAATGTTTCAATTTTGGAG--AGTTTCATAGGGTGTGCAAGAGAGGTACAGTC- 533
Db 73796 CTGACTATGTTTCAGTTTTCAGAGAAAGTCTGTGAGGTGCGAGAGAGGTATATTCT 73855

Qy 534 TTGTGTTTTGGTGAATAGTCTGTAATATCT-CTAGGTCCACTGTTGTTTATGACATCA 592
Db 73856 TTTGTGTTGGCGAGATGTTCTGTAATATCTGTAGGTCCATTTGTTTATGATATCA 73915

Qy 593 GTTAGCTCCAGATTCTCTGTTTCGTT-----TTTGTGTAGATGACCTTAACGTGTGG 646
Db 73916 GTGAGCTCCAGATTCTCTGTTAGTTTGTAGTTTGTGTAGTACCTGTCATNGGG 73975

Qy 647 AGAGAAATGGGTATTGAAGTAGGACCACTATCTGTGTGTA--GGTCAATATGTAATTAG 705
Db 73976 CAAGAAATGGGTATTGGAGTCTCTCACTCTCAGTGTGTGAGGGTCAGTACAAATTTAAG 74035

Qy 706 CTGTAGCTGCTGTTTATGAACTTGGGTGCAATCTGTTGTTGGTGCATAGACATTAAG 765
Db 74036 GTGCAGTAGTCTCTTTTACAAACCCGAGTCTTTTGTGTTGGTGATACATTTAAG 74095

Qy 766 AATTCGAATGCTCTCTGTTGGAATTTCTTTGATGCCCTATGTATATTCCTCCCAATCT 825
Db 74096 AGTTGCAATGCTCTCGGTGGACTTTCCCTT--TGAGTAAGTAGTGCTTCCCTCTTT 74153

Qy 826 CATCTGCTTAGTTTGGGTTTAAGTCTATTAGTCAGATATTAAATGACTGTATCGGCTT 885
Db 74154 CTCTGATATGTTTATTGTAAGTCTGTTTGTGCAGACAGAAATTAGCTTCCCGGCTT 74213

Qy 886 GCTTCTTAGGGCAATTTGCTTAGAATATCTTTTCCATCTTTTACTCTAAGGTGATGTCT 945
Db 74214 GCTTCTTAGGTCTTTTACTCAATTTGGTTTTTTTATCTTTTACCCTGAGATATCC 74273

Qy 946 ATCCAGTAGGTG--TCTTTTTGGAGCAGCAGTAGGATAGGATCTGTTTTCATATC 1003
Db 74274 ATTCCTGATGTGAGGTGTATTTTGGATTTCAGCAGAGGATGTAT-----74319

Qy 1004 CATCTGTATCCAGTATCTTTTCTAGAGAAATTAAGATCATTCATGTCATTGATGTGA 1063
Db 74320 TATTCGTGTTTCAAAATCAATCTGTTATTTTAAATCGGGGAATTAGTCCATGATGTGA 74379

Qy 1064 GAAATATCAATGAGCAGGTTTGTGGATTTCTGTTATCTTGCACTTTGTGAAGTGTGTG 1123
Db 74380 GAGTCATAAATGAACAGTGGTGTGTTGGGTTTTGTTATTTTGTAGTGGGTAG-----74431
```

```
Qy 1124 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1183
Db 74432 -----74431

Qy 1184 TCTGTGTCTGTGTGTGTGTGTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1243
Db 74432 -----ATTCTTTCATTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 74471

Qy 1244 TTATTCATATTTCTTGAAGTGGGTAACATCTTAGATTAAGTTTCTCTCTAG-CCT 1302
Db 74472 TTTCTTGTGTGTCTTACGTGGTTAACTCTGCACATTACAGATTCTCTCTCTAGCCCT 74531

Qy 1303 TCTTTAGGTCTGCATTTGAAGATAGATATCTTCTACATCTGATTTTATCTTAGAATGTCT 1362
Db 74532 TTGATAGAGCTGAGATTGTAGATAGATATTGCTTAAATTTGATTTTATCAT--GGAA 74589

Qy 1363 TTTCTTCTCCAACTATTGTGCACAGAAAGTTTCTTAAAGTGCAGTAGCTGGCT--GACAT 1421
Db 74590 ATATTTCTCCATCTATTTTGAATGAAATTTTGTGGGTATAGTAGCTGGATGGACAC 74649

Qy 1422 CTGTAGTCTCTGGAGTCTGTAGACATCTGTGCAGGCTTCTTACATTTTGAGTTTCT 1481
Db 74650 CTGTGCTCTGTTGGAGTTTGTAGACATCTGTCCAGGACTTTCTGGCTTCTAGTGTCTCC 74709

Qy 1482 ATTGGAAAAGTCAAGTGTAAATCTTAATACATCTGCCCTTTATATGTTAAATGGTCTTTT 1541
Db 74710 ATT-GAAAAGTCAAGTCTAAATCTAATAGATCTGCC-TTATATGTTATTAATAGCTT-- 74765

Qy 1542 CCCTTGCACTTTTAAATATCTTCTTCTCTATATCTTCTAGTGAATTTGATTTATGTC 1601
Db 74766 -CCTTGCACTTTTAAATATCTTCTTCTCTATATTTAGTGTGTTTGGAGTACCATGT 74824

Qy 1602 ACTGTGG---GAGTTTCTTCCGTCCTCAATCTATTTGGTGTGTTTGTATGCTCTGT 1657
Db 74825 GCTGTGTGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 74884

Qy 1658 ACCTTGATAGGACTCTCTTCTCAAGGTTAGGAAATTTTCTTTTGGTTTCTTTGTTTGA 1717
Db 74885 ACCTTGATAGGACCCCTGTTCTTATATAGTAGGAACTTTTCTTCTACGATTTTGTGGA 74944

Qy 1718 ATATTTTCCCTGCTTTTGACCT--GCCTTCTCCCTTCCCTCTATTC-----1763
Db 74945 ATATTTCTGTGTCTTGTAGACCTGGGTTTCTCCCTTCTCTATTCCTACTATTTGTAGA 75004

Qy 1764 TTTGGTTTTGTAGTAGTCTCTGGCTTCTCGATCTTCTGATCTTTTATGCCCTGGATTTT 1823
Db 75005 TTTGGTCTGTTCACAGTCTTTTGGTTTATGATATTTTGGCTGGATTTCTTTCATGT 75064

Qy 1824 TTAACATTTTCTTGACCAAGGTATCCATTTCTTATCTTGTCTTCACTGCTGAGATT 1883
Db 75065 TTCATATTGACTTTGACCAAGGTATCCATTTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCT 75124

Qy 1884 CTCTCTTCTATCTTGT--ATTCTGTCAAGTGGCTTGTCTCTGAGGTTTCTGTTGGGTT 1942
Db 75125 CTCTCTTCCCTCTCTTGTGTTTCTGTGTGTAAGGCTCACCTGAGATTTCTGTTCAAGT 75184

Qy 1943 CTTAATTTTTCATTTCCAGATTTCCTCAGTTTGGGTTTGTGTTTATTAATTTCTATTC 2002
Db 75185 TCAGAAATTTTTCATTTCCATTTTCTCCAGTTTGGGTTTCTTATTAATTTCAATTTCC 75244

Qy 2003 ACTTTCAGGCTCGAAATGTTTACTCATTT---TTCTCTCCAGTATTTACATTTTCATAG 2059
Db 75245 ACTTTCAGGCTTTGAACTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 75304

Qy 2060 GTTCTTTTAAATGATTTATTTCA--TTTCTCTTCAAGGACTTTTATGAATTTCAATAAATG 2118
Db 75305 ATTAGTAAATGATTTATTAATTTTCTTATGGGCTCTAATATATTTATTAAGGC 75364

Qy 2119 TATGTTAAGGCTCTGCTTGTCTTCAAGTATGTTG-----CATCTCAGGCT 2169
Db 75365 TATTTGGAGTCTTGTCTTGTGACTGAGTGTATTGTGATCCTCATGTTCTAATAGT 75424
```

QY 2170 ATTGTAATAGGTTTACAGGACATATTGCTCGGTGTTGTTATGTTCTGTTGTTTGCCT 2229
 Db 75425 AGGGATGTTGGCTCTAGTAGGGGCAATTGTTCCAGCTTTATTGACTGTGTTTAATT 75484

QY 2230 TTGGCATATAGACGCTCAGTTTGGGATGATTGTAATTCTAGGTCGTCGAT 2279
 Db 75485 CTGATGCTAGGCATCTGGGTTTGGGATGATTGTTAGTTCTAGTAGTCGAT 75534

RESULT 12
 AC124411 177414 bp DNA linear ROD 13-NOV-2003
 LOCUS Mus musculus BAC clone RP24-267C20 from chromosome 13, complete
 DEFINITION sequence.
 AC124411
 VERSION AC124411.5 GI:29124267
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 177414)
 VanBrunt, A., Bielicki, L., Van Brunt, A. and Kozlowski, A.
 The sequence of Mus musculus BAC clone RP24-267C20
 Unpublished (2001)
 2 (bases 1 to 177414)
 Wilson, R.
 Sequencing of Mus musculus
 Unpublished (2001)
 3 (bases 1 to 177414)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 4 (bases 1 to 177414)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (06-FEB-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 5 (bases 1 to 177414)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (19-MAR-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 6 (bases 1 to 177414)
 Wilson, R.
 Direct Submission
 Submitted (13-NOV-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Mar 19, 2003 this sequence version replaced gi:28261652.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics

 Center project name: M_BB0267C20

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RP24-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES
 source
 1..177414
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="13"
 /map="13"
 /clone="RP24-267C20"
 /clone_lib="RPC1-24"
 1..48
 /rpt_family="Alu"
 147..249
 /rpt_family="L2"
 391..509
 /rpt_family="L1"
 646..828
 /rpt_family="B2"
 1207..1634
 /rpt_family="L1"
 2321..3214
 /rpt_family="L1"
 2420..2803
 /note="Sequence derived from PCR product of project DNA."
 3391..3798
 /rpt_family="MaLR"
 3799..4204
 /rpt_family="MaLR"
 4819..4916
 /rpt_family="L2"
 4925..5391
 /rpt_family="RMR17C"
 5049..5393
 /rpt_family="ERVK"
 6462..6779
 /rpt_family="MaLR"
 7374..7583
 /rpt_family="B4"
 7877..8013
 /rpt_family="MaLR"
 8215..8336
 /rpt_family="MaLR"
 8696..8914
 /rpt_family="L2"
 9013..9297
 /rpt_family="MaLR"
 9882..10179
 /rpt_family="B4"
 10601..10682
 /rpt_family="L1"
 11927..12038
 /rpt_family="Alu"
 13322..13594
 /rpt_family="B4"
 14823..15002
 /rpt_family="MaLR"
 15218..15612
 /rpt_family="MaLR"
 15613..16707
 /rpt_family="MaLR"
 16708..17102
 /rpt_family="MaLR"

Query Match 33.0%; Score 752.002; DB 1; Length 177414;
 Best Local Similarity 66.6%; Pred. No. 0.071;
 Matches 1340; Conservative 0; Mismatches 515; Indels 158; Gaps 20;
 Qy 294 TATTCTTCATTCTATCTTGCTCATTTTTAACTCAGTAGTGGTGTGGTTTCCAT 353
 Db 151747 TCCTTCTAAATATCATCTTGACTCAATTTTTCATTCAGTAGTTCAGTGTGTCAT 151806

QY 1424 GTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCCCTTCTTACATTTTGAGTTCTAT 1483
 Db 152778 GTAGTCTCTTGTAGTCTGTGAGAACATCTGTCCAGGCTCTTCTGGCTTTTAGAGTCTCCAG 152837
 QY 1484 TGGAAAGTCAAGTGAATCTTATAACATCTGCCTTATATGTTAAATGGTCTTTTTC 1543
 Db 152838 T-GAAGTGGGTATAATCTATCAGGTC-GCCTTTATATGTACTTTTC-TTTTCCC 152895
 QY 1544 CTTCATCTTTTAATATCTTTCTTGTCTATATCTTTAGTGAATTCATTAATATGAC 1603
 Db 152896 CTTCAGCTTAAATATCTCTTATTTTGTATGTT-----AGTATTTATAGGTG 152949
 QY 1604 TGTGGGAGTCTTTTCCGGTCCAACTATTTGGTGTATTTGTATGCTTTCTGTACCTTG 1663
 Db 152950 TGAAGGACATTTCTTTCTGTGTAGTCTACTTTGTGTCTATATGCTTCTGTACCTTG 153009
 QY 1664 ATAGGCATCTCTTCTCAAGGTTAGGAATTTTCTTTTGTGTTTCTTCTGAAATATTT 1723
 Db 153010 ATAGATGCTCTTT-----TGGTTAGGAAGTTTCTTCTATGATTTGT-TGAAATATTT 153065
 QY 1724 TCCCTGCTTTTGACCTGCTTCTTCCCTTCTCTA-----TCCCTTGGTTTTCGA 1776
 Db 153066 TCTGTGCTGTGTTTCTTCTTCTCTCAATTTATATTAATTTAGATTTTCTTCTG 153125
 QY 1777 TAGTCTCTGCTTCTCTGGATTTTATGCTGGATTTATTTAGACTTACATTTCTT 1836
 Db 153126 TACTGTCCAAGATTCCTTGGATTTTATGCGAGATTTTATAGACTTAACATTTTACT 153185
 QY 1837 TGACCAAGATATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATTTCTCTTCTATCT 1896
 Db 153186 TGACTAAGTAGTCTTCTTCTTCTATCAATGCTTCAATGCTGAGACTCTCTTCTATCT 153245
 QY 1897 CTGTATCTGTGAGGCTGTCTGAGGTTCTCT-GTTGGGTCTTAAATTTTCA 1955
 Db 153246 CTGTATCTGTGAGGTTGTGCTGAGGTTCTCTGAGGTTCTCTTAAAGTCTTCTA 153305
 QY 1956 TTCCAGATTTCTTCAAGTGGTGTGTTTATTAATCTTCAATTTCACTTCAAGTCT 2015
 Db 153306 -----AACCCTTAGTCTGGTCTTCTTATGATTTCTTCCAGTTTCTGCTCT 153356
 QY 2016 GAAATGTTTACTCATTTCTCCAGTATTTACATTTTCAATTTTCAAGTTCTTTTAA 2075
 Db 153357 CAAGCATATTTTCA-TCTCTGTTTGTGCTGTATGTTTCTCAGATTTCTTTAAGGGGA 153415
 QY 2076 TATTCATTTCTCTTCAAGGACCTTTTATGATTCATTAATGATGTTAGGTCCTTGC 2135
 Db 153416 TATTCATTTCTCTTCAAGGACCTTCTATCATATTCATTAAGGTTATTTAAGGCCAT 153475
 QY 2136 CTGTGCTTCACTATGTTGC-ATTCTCAGGCGCTTATTTGTAATAGGTT-----TTTA 2186
 Db 153476 CTGTGCTTCACTATGTTTACAATCTCAGGAAATGCTGTGTAAGGTTTACTGAGCTTA 153535
 QY 2187 GCAGGACATATGTCCTGTTGTTTATGCTGTGTTTCTTGGCATATAGACGCT 2246
 Db 153536 GTGAGACATATGTCCTGCTGCTGTTGTTGATTTGTTATGTTATGCTAGACCTAGG 153595
 QY 2247 GAGTTTGGATGATTTGTAATTTAGGCTGCTGAT 2279
 Db 153596 GCATGTGGGAGACGTGTAATTTAGTCTGAT 153628

RESULT 13
 AL672047

LOCUS AL672047 184474 bp DNA linear ROD 18-SEP-2002
 DEFINITION Mouse DNA sequence from clone RP23-316B4 on chromosome X, complete sequence.

ACCESSION AL672047.9 GI:23093223

VERSION AL672047.9

KEYWORDS Mus musculus (house mouse)

SOURCE HTG.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 184474)
 Leongamornlert, D.
 Direct Submission
 Submitted (11-SEP-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On Sep 17, 2002 this sequence version replaced GI:22265406.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em., EMBL; Swi.,
 SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
 database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-316B4 is
 from the RPI-23 Mouse PAC library
 constructed by the group of Pieter de Jong.
 For further details see http://www.chori.org/bacpac/home.htm
 VECTOR: pBACe3.6.

FEATURES

source

1. 184474
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosomes="X"
 /clone="RP23-316B4"
 /clone_lib="RP23-23"

Query Match 32.0%; Score 729.201; DB 1; Length 184474;
 Best Local Similarity 68.7%; Pred. No. 0.097;
 Matches 1295; Conservative 0; Mismatches 508; Indels 82; Gaps 22;

QY 432 AGPATTTTCAATTTGCTTTTATCTGTCGAGACTTGTCTTTGTAATATGATTTCA 491
 Db 178221 AGATGTTTTCAATTTTCTTTCATCTGTAGAGATTTGCTTTTGGAGAATAATGGTCA 178280
 QY 492 ATTTTGGAG-AGTTTCATAGGTCGTGACAGAGGTACAGCTTTTGTGTTGGTGA 549
 Db 178281 ATTTTGGAGAGTTCATGAGGTCGTAATTAATGGAGGCA-TGTTTTATGTTTGGTCGG 178339
 QY 550 ATAGTCTGTAATATCT-CTAGGTCACCTTGGTTTATGACATCAGTTAGCTCCAGCATTT 608
 Db 178340 ATGTTCTGTAATATCTGTTAAGTCCATTTGATTTATATAATCAGTTAACTCCAGCATTT 178399
 QY 609 CTCTGTTTCTGTTTTTGTGAGATGACCTAACTGTTGGAGAAATGGGTATTGAAGTAG 668
 Db 178400 TTCTGTTTTGTTTTGGCTGA-TCAATCCATCTGTGTACTAGAGAGGTTATGAAGTTT 178458
 QY 669 CCCACTATCTGTGTGAGGTCATATGATTTTAGCTGTAGCTGTGCTGTTTTTATGA 728
 Db 178459 CCCACTATCAGTGTGTGAGAGCCATATGATGATTTAAGCTAATAGTATTTCTTTTATAA 178518
 QY 729 ACTTGGGTGACATTTGTTTGGTGCATAGACATTAAGAAATTCATGCTCTCTTGGTGA 788
 Db 178519 ACTTGGGTGCTCTTGCATTTGGTGCATAGATGTTTAAATTAAGTCTCTTGTATGA 178578
 QY 789 TT-TTCTTTTGATGCTATGATGATTTCTTCCCAATCTCATCTGCTTAGT-ITTGGG 843

Ds	178579	TTATTCCTTTTATGAGTATGAGTTTCTCTTCTTATCTTATGATGAGGGGTTTTT	178638
Qy	844	TTTAAGTCTATTAGTCAGATATAAATGACGTATCGGCTTGCTTCTTAGGGCCATTGG	903
Ds	178639	TGAACCTGTTTATCTGATATATAATATCTACCAACTTGTCTTAGAGTCCATTGG	178698
Qy	904	CTTGAATATCTTTTCCATCCTTTTACCTCTAAGGTGATGTATCC--ATGGTAGGTGT	961
Ds	178699	CTTGAACCTACTTTTCCCAACATTTACCTCAGGTGATGTATCCTTACGTTAGGATGT	178758
Qy	962	CTTTTTCGGA--TGCAGCAGTAGGATGATCTGTTTTCATATCCATCTGTATCCAGTA	1020
Ds	178759	ATTCTTUGATGTCCAGAGAGGGGATCTGTTTTCATATCCATCTCTTCTGCTT	178818
Qy	1021	TCCTTTTCT-----AGAGAAATTAAGATCATATGATCATGATTTGAGAATT	1068
Ds	178819	GTCCTTTTATTGGGAATTCAGACCACTGAAGTTGAACAGCAGTCTTGTGATTAATT	178878
Qy	1069	ATCAATGACAGTGTGTGATCTTGTATCTTGCACATCTGGAAGTGTGTGTGT	1128
Ds	178879	GCTATTTTGT	178938
Qy	1129	GT	1188
Ds	178939	TGGTAGTAGT	178996
Qy	1189	GCTTGT	1248
Ds	178997	-----TCCTTTGATTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	179035
Qy	1249	CATATTTTCTTGAATGCGGTAAACAT--CTTTAGATTTGAAGTTTTCCTCTAG--CTTCT	1305
Ds	179036	TGGGTTTTCATGGGTAGTATTAACCTACTTATTTGGATTTTTCCTTCTAGTATCTTT	179095
Qy	1306	TTAGGTCCTGCAATTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAATCTTTC	1365
Ds	179096	GTAAGCTGGATTTATACACAGATATGCTTA-----AATTTTTCATGGAACATCTTAG	179150
Qy	1366	TTTCTCBACTATTGTGACAGAAAGTTTTCATAGTCAGTAGTCTGGCCCTGACATCTGT	1425
Ds	179151	TTTCTACATCTATTATGATGAAAGTCTGTGGGTATAGTAGTCTGGACTGGCAATTTAT	179210
Qy	1426	AGTCTCTTGGAGTCTGTAGCACATCTGTGACAGGCTTCTTACATTTTGAATTTCTATTG	1485
Ds	179211	GCTCTCTTAAAGTCTACAGCACTGTGCCAGTCTTCTGCTTTTAAATTTCCATT--	179269
Qy	1486	GAAAGTCAGGTGTAATTTCTAATACATCTGCCTTTATAGTTAAATGCTCTTTTCCCT	1545
Ds	179270	GAGATCTCAGGCATAAATCTAATAGTCTGCTTATATGTCACTTGTATC--TTTACCCCT	179328
Qy	1546	TGCATCTTTTAAATTTCTTTCTTTCTTATCTATCTTTTAGTATTTGATTTATGCACTG	1605
Ds	179329	TGCACCTTTTAAATTTCTTTCTTTCTTATGTTAGTTAAATTTTATTTAGTATA	179384
Qy	1606	TGGGAGTTTCTTTCCGGTCCAATCTATTTGGTGTGTTTGTATGCTCTTCTGATCTTGAT	1665
Ds	179385	ATGGGAATTTCTTTCTTGTCTAAT--ATTGATGTTCTGTATGCTTCTTGTACCTGGAT	179441
Qy	1666	AGGCATCTCTTCAAGGTTAGGAATTTTCTTTTGTGTTTCTTGTGAAATTTTTC	1725
Ds	179442	AGGCATGCTCTTCTTTAGGTAGCAAGCTTCTTATGTAATTTTGTGAAATTTTTC	179501
Qy	1726	CCTGCTTTGACCTGCTCTTCTTCCCTCTATCTCTTGTGTTTGTGATGCTCTC	1785
Ds	179502	TGTGCTTTGACCTGGGTTTCTTCTTGTGTTATCTTCTAGATTTGTTTGTGCTGATCA	179561
Qy	1786	TGGCTTCTGGAATTTTATGCTGGATATTTTAGACTTAACATTTTCTTTTGACCAAGG	1845
Ds	179562	AGGCTTCTGATATTTTGTGCCAAGAGCTCTTTAGATTTTATCATCTTCTTCTGCTGGT	179621
Qy	1846	TATCCATTTCTTCTATCTGCTCTTCACTGCTGAGATCTCTCTTCTATCTCTTGTATTC	1905
Ds	179622	TATCCATTTCTTTTATCATGTCTCTTAATGCTGGGAGTTTCTTTTATCTCTTGTATTC	179681
Qy	1906	TGTCAGTGAGGCTTGTCTCTGAGGTTTCTCTGT--TGGGTTCTTAATTTTTCATTTCCAGAT	1964
Ds	179682	TGCTGAGGAGCTTGGCTCTGAGTTTCTGTCCTCTTTTACATTTTTCATTTCCAGAT	179741
Qy	1965	TTCCCTTCAAGTTGGGTTTGTGTTTATTAATCTATTTCCATTTCCAGTCTGAAATGTTT	2024
Ds	179742	TTTCCCAAGTTTCCATTTTCTTATTAATTTTATTTTCACTTTTCACTTTGAACAGTTT	179801
Qy	2025	TACTCATTTTCTCCCAAGTATTTACATTTTCAAGGTTTCTTAAATGGAATTTATTTCAATT	2084
Ds	179802	TATCTA--TTTCTTCTACTATTTGTGTTTTCATAGATTTCTTTTCAAGTATTTATTTCAAT-	179859
Qy	2085	CCTCTTCAAGGACCTTTTATGAATTCATAAATGTAATGTAAGTCTTCCCTTGTGCTT	2144
Ds	179860	-----TTTAAGGATCCCTATCATATTTCAAAAGCTATTTTAAATATTTTCTTGTGCTT	179915
Qy	2145	CAGCTATGTTGCATTT-----CTCAGGCTATTTGAATAGGTTTTCAGCAGGACA	2195
Ds	179916	CATTAATGTTAATGCTCAGGACCTGCTGTGGTGGAGTTGTTGGGTTTTTATGAGAGACA	179975
Qy	2196	TATGTCCTGGTTTGTGTTTATGCTGTG--TTTTTGTCTTGGCATATAGACGCTGAGTTTG	2254
Ds	179976	TACTGCTCAGCTGTTTATTAATGTTGTTTTCATCTGCTTACACTGCTGCTAAGATCTGCTTGG	180035
Qy	2255	GATGATTTGAATTTCTAGGTGCTGAT 2279	
Ds	180036	GAAGATTGTAATTTCTAGGTGTTGAT 180060	
RESULT 14			
AC093354			
LOCUS			
DEFINITION Mus musculus clone 204031 bp DNA linear HTG 20-AUG-2002			
AC093354			
VERSION AC093354.2 GI:22325146			
KEYWORDS HTG; HTGS PHAS1; HTGS DRAFT; HTGS_FULLTOP.			
SOURCE Mus musculus (house mouse)			
ORGANISM Mus musculus			
REFERENCE 1 (bases 1 to 204031)			
AUTHORS Birren, B., Nusbaum, C. and Lander, E.			
TITLE Mus musculus, clone RP23-20D11			
JOURNAL Unpublished			
AUTHORS 2 (bases 1 to 204031)			
TITLE Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hago, B., Headford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Menues, L., Michova, T., Miengo, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phurkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viell, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.			
JOURNAL Direct Submission			
TITLE Submitted (20-AUG-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			

REFERENCE 3 (bases 1 to 204031)
 AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cooke,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fero,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horcon,B., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Melidrim,J., Meneus,L., Mihova,T., Mienga,V.,
 Murphy,T., Navlor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Testfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliiev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE Direct Submission
 JOURNAL Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 20, 2002 this sequence version replaced gi:15213899.
 COMMENT All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genom Center
 Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIGR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: LI3634

Center clone name: 20.D.11

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 203365 bases at least Q40

Consensus quality: 203606 bases at least Q30

Consensus quality: 203724 bases at least Q20

Insert size: 194000; agarose-fp

Insert size: 203831; sum-of-contigs

Quality coverage: 14.1 in Q20 bases; agarose-fp

Quality coverage: 13.4 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

* 1 121934: contig of 121934 bp in length

* 121935 122034: Gap of 100 bp

* 122035 142040: contig of 20006 bp in length

* 142041 142140: Gap of 100 bp

* 142141 204031: contig of 61891 bp in length.

Location/Qualifiers

1..204031

/organism="Mus musculus"

/mol_type="Genomic DNA"

/db_xref="taxon:10090"

/clone="RP23-20D11"

/clone_lib="RPCI-23 Female Mouse BAC"

1..121934

/note="assembly_fragment"

clone_end:SP6

vector_side:left

122035..142040

/note="assembly_fragment"

142141..204031

/note="assembly_fragment
 clone_end:r7
 vector_side:right"

Query Match 32.0%; Score 725.201; DB 1; Length 204031;

Best Local Similarity 68.7%; Pred. No. 0.1;

Matches 1295; Conservative 0; Mismatches 508; Indels 82; Gaps 22;

QY	432	AGTATTATTTCAATTCCTTTTATCTGTCGACACTTCTTTTGTGGAATATGATTCACA	491
DB	177336	AGATGTTTTTCAATTTCTTTTCATCTGTAGACATTTGCTTGTGAAGAATATGGGTCA	177395
QY	492	ATTTTGGAG--AGTTTCATAGGTCGTGACAAAGAGTACAGTCTTTGTGTTTGGTCAA	549
DB	177396	ATTTGGAGAAAGTTCATGAGTGCTAAATGGAGGCA-TGTTTTATGTTTGGTCGG	177454
QY	550	ATAGTCTGTAATAATCT-CTAGGTCCACTTGGTTTATGACATCAGTTAGCTCCAGATTT	608
DB	177455	ATGTTCTGTAATAATCTGTTAAAGTCCATTTGATTTTATATATATCAGTTAACTCCAGATTT	177514
QY	609	CTCTGTTTCGTTTTTGTGAGATGACCTAACCTGTTGGAGAAATGGGTATTGAAGTAG	668
DB	177515	TTCTGTTTTGTTTTTGCCTGA-TCAATCCATCTGTTGACTAGTAGGGTATGAGTTT	177573
QY	669	CCCACTATCTGTGTGAGGTCAAATATGTAATTTAGTGTGTAGCTGTGCTTTTATGA	728
DB	177574	CCCACTATCAGTGTGTGAGAGCAATATGTGATTTAAAGCTTAATAGTATTTCTTTATAA	177633
QY	729	ACTTGGTGACATTTGTTTGGTCATAGACATTAAGATTCGAATGCTCTGTCGGA	788
DB	177634	ACTTGGTGCTCTTTCGATTTGTCATAGATGTTAAATTTGAAAGTCTTCTTGATGA	177693
QY	789	TT-TTCCTTTGATGCTCTATGATATTCTTCCCAATCTCATCTGCTTAGT-----TTTGGG	843
DB	177694	TTATTCCTTTTATGAGTAGTAGTTTCTTCTCTTATCTTATGATTAGTGGGTTTTT	177753
QY	844	TTTAAGTCTATTAGTCAGATATTAATAAGTGTATCGGCTTGTCTTAGGGCCATTGG	903
DB	177754	TGAACCTCTCTTTTATCTGATATTAATAATCTACACCAACTGTCTTCTAAGGTCCATTGG	177813
QY	904	CTTAGAATATCTTTTCCATCTTCTTACTCTAAGGTGATGTCTATCC--ATGTTAGTGTGT	961
DB	177814	CTTGGAAAGTACTTTTCCCAACATTTACCTGAGGTGATGTCTATCTTACGTTAGATGT	177873
QY	962	CTTTTTTGA-TGCAGCAGTAGATGGAFTTGTGTTTTCATATCCATTCGTTTACCAGTA	1020
DB	177874	ATTTCTTGATTTGCCAGAGAGGATCTTGTGTTTTCATCCATCTCTTTGTCTGT	177933
QY	1021	TCCTTTTCT-----AGAGAAATAGATCAATGATGATGATTTGAGAAAT	1068
DB	177934	GTCTTTTATTTATGGGAAATGAGACAGGAAAGTTGAACAACAGAGTCTTGTGTATAAT	177993
QY	1069	ATCAATGACAGTGTTTGTGGATTCCTTGTATCTTGTCACTTGTGAAGTGTGTGTGTGT	1128
DB	177994	GCTATTTTGTGTGTGTTTGTGCTGCTGTTGCTGTTGTTTCTTGTGTTCTTGTGTGG	178053
QY	1129	GT	1188
DB	178054	TGCTAGTAGT	178111
QY	1189	GTCTGT	1248
DB	178112	-----TCTTTGATTTGCTGGTACTGGAATTTTACTTATCTCT	178150
QY	1249	CATATTTTCTGAATGCGGTAAACAT-CTTTAGATGAAAGTTTTTCTCTCTAG--CCTTCT	1305
DB	178151	TGGGTTTTTCATGGGTGTAGTTAACTTACTTTATTTTGGATTTTTTCTCTAGTATCTTTT	178210
QY	1306	TTAGGTCTCATTTTGAAGATAGATATTCTTTACATCTGATTTTATCTTAGAATGTCTTTC	1365
DB	178211	GTAAGGCTGGATTTATAGACAGATATTCCTTA-----AATTTTTCAAGAACATCTTAG	178265
QY	1366	TTTCTCCAACTATTGTGACAGAAAGTTTTTCTTAAGTGCAGTAGTCTGGCCTGACATCTGT	1425

FEATURES
 source

misc_feature

misc_feature

misc_feature

Db	178266	TTTTTACATCATTTATGATGAAAGCTCTTGGGTATAGTAGTCGGACTGGCATTTAT	178326
Qy	1426	AGTCTTTTGGAGTCGTAGCACATCTGTGCAGGGCTTCTTACATTTTGAGTTTCTATG	1485
Db	178326	GGTCTCTTAAAGCTCTACAGCACATCTGTCCAGTGTCTTGGCTTTTAAATTTCCATT-	178384
Qy	1486	GAAAGTCAGGTGAATTTCTAATAACATCTGCCTTTATATGTTAAATGSPCTTTTTTCCT	1545
Db	178385	GAGATCTCAGGCATAATTTCTAATAGTTCTGCCCTTTATATGTCACITGATC-TTTTACCC	178443
Qy	1546	TGCATCTTTTAATATCTTTCTTTCTGTCTATACITTTTAGTGATTTTGATTTATGACATG	1605
Db	178444	TGCACCTTTTAAATATCTTTCCCTCTCTGTAAGTTAAAGTT- ---TTATATAGTATA	178499
Qy	1606	TGGGAGTTCTCTTTTCCGGTCCAACTATTTTGGTGTTTGTATGCTCTCTTGACCTTGAT	1665
Db	178500	ATGGAAATCTCTTTCTTTGTCTAAT--ATTGATGTCTGTATGCTCTCTTGACCTGGAT	178556
Qy	1666	AGGCATCTCTTTCTCAAGGTAGGAAATTTTCTTTTGGTTTCTTGAAATAATTTTC	1725
Db	178557	AGGCATGTCCCTTTTAGTTAGCAAGCTTTCTATGTAATTTTGTGAAATAATTTTC	178616
Qy	1726	CCTGCTTTTGACCTGCCCTTCTCCCTCTCTATTCCTTTGGTTTGTSCATAGTGTCTC	1785
Db	178617	TGTGCTTTTGACCTGGGTTTCTTCTCTGTATTCTCTAGATTTGTTTTTGTGGTGTACA	178676
Qy	1786	TGGCTTCTCGAATGTTTATGCTCGAATATTTTAGACTTAACATTTTCTTTGACCAAGG	1845
Db	178677	AGGCTTCATGGATATTTTGTGCCAAGAGCTCTTTAGATTTATCACTTTCCITGACTGGTG	178736
Qy	1846	TATCCATTTCTCTATCTTGCTTCCACGTCCGTAGATCTCTCTCTATCTCTTCTGATTC	1905
Db	178737	TATCCATTTCTTTATCATGTCCTTAATGCCCTGGAGTTTTTCTTTTATCTCTTGTATTC	178796
Qy	1906	TGTCAGTGAGGCTTGCTCTCAGGTTCCCTGT-TGGGTTCTTAAATTTTTTTCATTTCCAGAT	1964
Db	178797	TGCTGGTGAAGCTTGCCTCTGAGTTTCTGTCCCTCTTTTACATTTTCAITTTCCAGAT	178856
Qy	1965	TTCTTTCAGTTTGGGTTTGTTTATTAATCTTAATTTCCACTTTTCCAGTCTCTGAAATGTTT	2024
Db	178857	TTTCCCGAGTTTCCAAITTTCTTTATTTATTTTATTTTCAGCTTTTCAGTCTTGAACAGTTT	178916
Qy	2025	TACTCATTTTCTCCCCAGTATTACATTTTCATAGTTTCTTTTAATGGAATTTTATTCATTT	2084
Db	178917	TATTCA-TTTCTCTCTACTATTTTGTGTTTTTCAAGATTTCTTTTCAAGTATTTATTTCAIT-	178974
Qy	2085	CCTCTTCAAGGACCTTTTATGAATTCATAAAATGATGTTAAGGTCCTTGCCCTGTGCTT	2144
Db	178975	---TTTAAGGATCCCTATCATATTATAAAGGCTATTTTAAATAATTTTCTTGTGCTT	179030
Qy	2145	CAGCTATGTTGCATT-----CTCAGGGCCTATTGTAATAGGTTTTTACGAGGACA	2195
Db	179031	CAGTTATGTTGTAATGCTCAGGACCTGCTGTGGTGGAGTTGTGGGTTTTAGTGAGACA	179090
Qy	2196	TATTGCTCTGGTGTATTGTCGTG-TTTTTTGGCTTTGGCATATAGACGGCTGAGTTGG	2254
Db	179091	TACTGCTCTAGCTGTTATGATTGTGTTTTTACACTGGTGTCTAAGATCTTGGCTTGG	179150
Qy	2255	GATGATTTGAATTTCTAGGTGCTGAT	2279
Db	179151	GAAGATTTGAATTTCTAGGTGTTGAT	179175

RESULT 15					
AL806512/c					
LOCUS	Al806512	21131 bp	DNA	linear	ROD 07-NOV-2002
DEFINITION	Mouse DNA sequence from c-lone RP23-354A20 on chromosome 4, complete sequence.				
ACCESSION	AL806512				
VERSION	AL806512				
KEYWORDS	AL806512.9	GI:24527497			
SOURCE	HTG.				
	Mus musculus (house mouse)				

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus
1 (bases 1 to 211131)
Martin,S.
Direct Submission
Submitted (01-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk clone requests: clonerequest@sanger.ac.uk
On Nov 3, 2002 this sequence version replaced gi:23337500.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-35A20 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6.

```

FEATURES
source
LOCATION/Qualifiers
1. .211131
/organism="Mus musci
/mol_type="genomic
/db_xref="taxon:100
/chromosome="4"
/clone="RP23-354A20
/clone_lib="RPCI-23

```

Query Match	31.8%;	Score 725.001;	DB 1;	Length 211131;
Best Local Similarity	68.11%;	Pred. No. 0.11;		
Matches 1402;	Conservative 0;	Mismatches 480;	Indels 178;	Gaps 30;
Qy	290	CTTCTATTCTTGATTCTATCTGGGTCATTTAACTCAGTAGTGAGTGGTTGTG---	345	
Db	164325	CTGTCTCTCTTGATTCTGTCTTGGCCCAATTTCCATTTAGTAGCAAGTGTTCAGTCA	164266	
Qy	346	GTTCCTCAATGAAGTTCTGAAGTTCTTCGTGTGTGTGTGTGTGTGTATCTAGATTTTA	405	
Db	164265	GTTCCTATGAGTTAGTAAGCCTTCTGTATTCTCTGC--TGCCTGCTGATCCAACTTTA	164209	
Qy	406	AGCTGTGGTGGTCAGATAGGCACATAGTAGTATTATTCAATTCGTTTTTATCTGCTCGAGAC	465	
Db	164208	GTAAGTGGTGGTCAGATAGGATGACGGGTGCGGTTTCCATTAACCTGTATTGTGTGAGAA	164149	
Qy	466	TTGCTTTGTTTTGAAATATGATTCAATTTTTGGAG--AGTTTCATAGGGTGCTGCACAAGA	523	
Db	164148	TTGCTTTTGGCCCAAGTATGTGGTCAGTTTTTGGAGAAAGTATGTCGTGGTGCTAAAAAGA	164089	
Qy	524	AGGTACAGTC-TTTGTGTTTTGGTGAATAGTCGTGAAATATC-TCTAGGTCCACTTGGT	581	
Db	164088	AGATATATCTTTGTGTTTGGGTGAAATGTTCTGTAATACCTTCTAGGTCCCTTTGCT	164029	
Qy	582	TTATGACATCAGTTAGCTCCAGCATTTCTCTGTTCGTTTTTGTTCGAGATGACCTAACT	641	

QY 1864 TGCTTCTACTGCTGAGATTCCTCTCTCTATCTCTGTTATCTCTGTCAGTGAGGCTTGCT 1923
 Db 18076 TGTCTTTAACAGCTGAGATTCCTCTCTCTATCTCTGTTATCTCTGTCAGTGAGGCTTGCT 18135
 QY 1924 CTGAGGTTCTCTGCTGCTT-CTTAATTTTTCATTTCCAGATTTCCCTTCAGTTGGGTTT 1982
 Db 18136 CTGAGGTTCTCTGCTGCTT-CTTAATTTTTCATTTCCAGATTTCCCTTCAGTTGGGTTT 18195
 QY 1983 TGTATTATTAATTTCTATTTCCATTTCCAGGCTGCTGAATTTTACTCATTTTCTCCACG 2042
 Db 18196 TCTTTATTGCTTTTATTTCCATTTCCAGGCTGCTGAATTTTACTCATTTTCTCCACG 18255
 QY 2043 TATTTAC-ATTTCATAGGTTCTTTAATGATTTATTTCCCTTCAGGACCTT 2100
 Db 18256 GTTCTCTGCTGCTGCTTCTGGAATTTCAAAGGATTTATTTCTCTTTAAGGCTTCT 18315
 QY 2101 TTATGAATTCATAAAATGATTTAAAGTCTCTGCTTCTGCTTCAGCTATGTTGCATTC 2160
 Db 18316 TATCACCCTTCAAGGTTGGTTTAAAGGTTATTTACTTGTCTTCAGCTGTTTGGATG 18375
 QY 2161 TCAGGCTC-----TATTGTAATAGGTTTATGACAGGACATATTCCTCGTTGTT 2211
 Db 18376 TCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18435
 QY 2212 ATTGCTGCTGTTTTCCTTGGCATATAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2271
 Db 18436 GTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 18495
 QY 2272 GTGCTGAT 2279
 Db 18496 GTACCAAT 18503

RESULT 17
 AC115852/c
 LOCUS AC115852 142529 bp DNA linear HTG 21-OCT-2002
 DEFINITION Mus musculus clone RP24-270L8, WORKING DRAFT SEQUENCE, 6 ordered
 pieces.
 ACCESSION AC115852
 VERSION AC115852.4 GI:24182195
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 142529)
 AUTHORS Birren, B., Nusbaum, C. and Lander, E.
 TITLE Mus musculus, clone RP24-270L8
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 142529)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhegater, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
 Lander, A., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, W., McEwan, P., McKernan, K., Meldrim, J., Meneus, L.,
 Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, D.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission

JOURNAL REFERENCE AUTHORS

Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 142529)
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhegater, B.,
 Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,
 Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T.,
 Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (21-OCT-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Oct 21, 2002 this sequence version replaced GI:21622796.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE JOURNAL COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L24684
 Center clone name: 270_L8
 ----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 141053 bases at least Q40
 Consensus quality: 141891 bases at least Q30
 Consensus quality: 141937 bases at least Q20
 Insert size: 145000; agarose-fp
 Insert size: 142029; sum-of-contigs
 Quality coverage: 9.5 in Q20 bases; agarose-fp
 Quality coverage: 9.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 6 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 1173: contig of 1173 bp in length
 * 1174 1273: gap of 100 bp
 * 1274 4401: contig of 3128 bp in length
 * 4402 4501: gap of 100 bp
 * 4502 99416: contig of 94915 bp in length
 * 99417 99517: gap of 100 bp
 * 99517 113174: contig of 13658 bp in length
 * 113175 113274: gap of 100 bp
 * 113275 140154: contig of 26880 bp in length
 * 140155 140254: gap of 100 bp
 * 140255 142529: contig of 2275 bp in length.

FEATURES source

1. 142529
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090".


```

Db      129264  TATCACCTTCATAGGTGGTTTAAAGTTATTTACTTGTTGTTTCAGCTGTGTGGAATG 129205
Qy      2161   TCAGGGCC-----TATTGTAATAGGTTTATAGCAGGACATATGCTGTGGTTGTT 2211
Db      129204  TCCAGAGCTTGCTGTAGTAGGATAGTTAAAGTTCCGGTGATACCAATGTTGCCAAGGCTATT 129145
Qy      2212  ATTGCTCTGTTTTCCTTTGGCATATAGACGCTGAGTTTGGGATCATTTGTAATTCAG 2271
Db      129144  GTTGATGTTGTTCTATGCTGGCTCTAAGTACCTGGATTGGGATATATAGGTTTAG 129085
Qy      2272  GTGCTGAT 2279
Db      129084  GTACCAAT 129077

RESULT 18
AC129304 153772 bp DNA linear ROD 27-NOV-2003
LOCUS Mus musculus BAC clone RP24-462H23 from chromosome 10, complete
DEFINITION
AC129304
AC129304.6 GI:30425602
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
Shah,N. and Meyer,R.
1 (bases 1 to 153772)
The sequence of Mus musculus BAC clone RP24-462H23
REFERENCE
AUTHORS Wilson,R.
TITLE Unpublished (2001)
JOURNAL 2 (bases 1 to 153772)
AUTHORS Wilson,R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE 3 (bases 1 to 153772)
JOURNAL Direct Submission
AUTHORS Submitted (28-JUL-2002) Genome Sequencing Center, 4444 Forest Park
TITLE Parkway, St. Louis, MO 63108, USA
JOURNAL 4 (bases 1 to 153772)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2003) Genome Sequencing Center, 4444 Forest Park
TITLE Parkway, St. Louis, MO 63108, USA
JOURNAL 5 (bases 1 to 153772)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (08-MAY-2003) Genome Sequencing Center, 4444 Forest Park
TITLE Parkway, St. Louis, MO 63108, USA
JOURNAL 6 (bases 1 to 153772)
AUTHORS Wilson,R.
TITLE Direct Submission
JOURNAL Submitted (27-NOV-2003) Department of Genetics, Washington
TITLE University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
JOURNAL On May 8, 2003 this sequence version replaced gi:28973954.
COMMENT
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
----- Summary Statistics
-----
Center project name: M_BB0462H23
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RP24-24 BAC library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC122018.

FEATURES	Location/Qualifiers
source	1..153772
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="10"
	/clone="RP24-462H23"
	/clone_lib="RP24-24"
repeat_region	326..459
	/rpt_family="Alu"
repeat_region	546..747
	/rpt_family="MIR"
repeat_region	1790..1966
	/rpt_family="MaLR"
repeat_region	1971..2070
	/rpt_family="B4"
repeat_region	2644..2685
	/rpt_family="B4"
repeat_region	2782..2964
	/rpt_family="MER1_type"
repeat_region	3483..3586
	/rpt_family="MIR"
repeat_region	3648..3768
	/rpt_family="LI"
repeat_region	3751..4429
	/rpt_family="LI"
repeat_region	5506..6144
	/rpt_family="LI"
repeat_region	6425..6643
	/rpt_family="ERVK"
repeat_region	6772..6818
	/rpt_family="ERVK"
repeat_region	6862..6953
	/rpt_family="Alu"
repeat_region	6958..7146
	/rpt_family="ERV1"
repeat_region	7886..8253
	/rpt_family="MaLR"
repeat_region	8533..8916
	/rpt_family="MaLR"
repeat_region	9222..9364
	/rpt_family="LI"
repeat_region	9365..9665
	/rpt_family="MaLR"
repeat_region	9666..9692
	/rpt_family="LI"
repeat_region	10041..10071
	/rpt_family="LI"
repeat_region	10086..10673
	/rpt_family="LI"
repeat_region	11040..11175


```

Db 21060 CTTTAGAGTGAATTAACCTTCCAGTGAATCTGTAGGGCTCATCTATTATATAATA 21119
Qy 1331 TTCCTTACATCTGATTTTATCTTAGAATGCTTCTTCTCCCACTATTGTGACAGAAG 1390
Db 21120 TTGCTTACATTTGGCTTTTATCATAAATGCTGTTTCTTCATCTATTGTGATGAAG 21179
Qy 1391 TTTTCTAAGTCAGTCTAGTCTGGCTGACATCTAGTCTCTCTGGAGTCTGTAGCACATC 1450
Db 21180 TTTTCTAGGTATAATAGTAGGCTGCACTGTGTCTCTTAACTTCGTAGCACAT 21239
Qy 1451 TGTGAGGCGCTTCTTACATTTTGAATTTTCTATTTGGAAGTCAAGTGTATCTTATAC 1510
Db 21240 TGTCTAGGCTCTTTTGTGTTTAGAGTCTC-----CATGAGAAGTCTGTAG 21287
Qy 1511 ATCTGCCCTTATATGTTAATTTGGTCTTTTTCCTTGCATCTTTTAACTTTCTTTCT 1570
Db 21288 CTCGCGCTTATATGTTACTTGG-CCCTTTTCCTTAAGGCTTTTAACTTTCTTTCT 21346
Qy 1571 TTCATATCTTTTAGTATTTGATTAATGACACTGTGGGGAGTCTCTTTCCGTCCTCAAT 1630
Db 21347 -TAITGTCAGTCAGTGTCTTGTGATAATATATGCTAAGGGGCTTTCTTTCTGTTCAAT 21405
Qy 1631 CTATTTGGTGTGTTGATGCTTCTTGTACCTGTAGGCACTCTTCTCAAGGTTAGGA 1690
Db 21406 CTCCTTGTGATTTCTAAATGC-----TACCCTTCTAGGCACTCTCTTCTTTAGTTAGGA 21459
Qy 1691 AATTTTCTCTTTTGGTCTTCTGAAATAATTTTCCCTGCTTTTGAACCTGCTTCTTCCC 1750
Db 21460 AATTTTCTCTATGATTTTGTGAAATAATTTCTATGACTTTTGAACCTGCTTCTCCC 21519
Qy 1751 C-----TTCTCTATCTCTTGGTGTGTTTGCATAGTCTCTGCTTCTGCTGATG 1799
Db 21520 CCCCCTCTCTATCTCTATTCCTTAGATTTAGCTTTGATCCCTCAATTTCTGCTGATG 21579
Qy 1800 TTTTATGCTGCTGATTTTATAGACTTAACATTTTCTTTTGAACCAAGGATCAATTTCTTCT 1859
Db 21580 TTTTATCATCAGGGTCTCTTAGAGTTACACTTTCTCTAACCAAGGATCAATTTTCT 21639
Qy 1860 ATCTTGTCTACCTGCTGAGATTTCTCTTCTATCTCTGATCTTGTATCTGTCTAGTGAGGCTT 1919
Db 21640 ACTGTGCTTTCAGTCTCAACATTTCTCTCTTCCATCTTCTGTTCTGTCTGTCTGAG- 21693
Qy 1920 GTCTCTGAGGTTCCGTGTTGTTCTTAATTTTTCATTTCCAGATTTCTCTTCACTTTGGG 1979
Db 21694 -----GTTTCTTACAGTGTCTTAATTTCTGTTTCAAGATTTCCCTCAGTTTGGG 21745
Qy 1980 TTTTGTATTATTAATTTCTATTTCCACTTTTCAGGCTCTGAAATGTTTTTACTCAATTTCTCTC 2039
Db 21746 TTTTGTATTATTAATTTCTATTTCCACTTTTCAGATCTGAACTGTTTCTATTTCA- 21804
Qy 2040 CAGTATTTACATTTTCATAGTCTTTTAAAGGATTTATTTCACTTTCTCTTCAAGGACCT 2099
Db 21805 CACTGTTTGTGTTTTCATAGATTTCTTTGTTAGATTTACTCACTTTCTCTTTTAAAGACA 21864
Qy 2100 TTTATGAA 2107
Db 21865 GTCAGGCA 21872

RESULT 19
AC122210
LOCUS
DEFINITION Mus musculus chromosome UNK clone RP23-96G7, WORKING DRAFT
SEQUENCE, 23 unordered pieces.
ACCESSION AC122210.2
VERSION AC122210.2
KEYWORDS HTG, HTGS PHASE1, HTGS_DRAFT, HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 271609)

```

```

McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 271609)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 271609)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Aug 25, 2002 this sequence version replaced gi:21105042.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0096G07
----- Summary Statistics -----
Sequencing vector: M13, 0%
Chemistry: Dye-primer ET; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 264453 bases at least Q40
Consensus quality: 266334 bases at least Q30
Consensus quality: 267385 bases at least Q20
Insert size: 234000; agarose-fp
Insert size: 268963; sum-of-contigs
Quality coverage: 15.12 in Q20 bases; agarose-fp
Quality coverage: 11.17 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1097: contig of 1097 bp in length
* 1098: contig of unknown length
* 1099: gap of unknown length
* 1198: contig of 56 bp in length
* 1254: gap of unknown length
* 1354: contig of 390 bp in length
* 1744: gap of unknown length
* 1844: contig of 1097 bp in length
* 2941: gap of unknown length
* 3041: contig of 1169 bp in length
* 4210: gap of unknown length
* 5593: contig of 1284 bp in length
* 5694: gap of unknown length
* 6894: contig of 1201 bp in length
* 6995: gap of unknown length
* 8293: contig of 1298 bp in length
* 8393: gap of unknown length
* 9864: contig of 1472 bp in length
* 9965: gap of unknown length
* 11951: contig of 1987 bp in length
* 12051: gap of unknown length
* 13051: contig of 1050 bp in length
* 13201: gap of unknown length
* 14843: contig of 1642 bp in length
* 14944: gap of unknown length
* 17129: contig of 2185 bp in length
* 17229: gap of unknown length
* 18762: contig of 1534 bp in length
* 18962: gap of unknown length

```

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

* 18863 25092: contig of 6230 bp in length
* 25093 25192: gap of unknown length
* 30277: contig of 5085 bp in length
* 30377: gap of unknown length
* 30378 47187: contig of 16810 bp in length
* 47188 47287: gap of unknown length
* 47288 63647: contig of 16360 bp in length
* 63648 63747: gap of unknown length
* 63748 83396: contig of 21649 bp in length
* 83397 85496: gap of unknown length
* 85497 112995: contig of 34499 bp in length
* 112996 147799: contig of 27704 bp in length
* 147800 147899: gap of unknown length
* 147900 194953: contig of 47054 bp in length
* 194954 271609: contig of 76556 bp in length.
* 195054 271609: contig of 76556 bp in length.

```

FEATURES

source

```

1. 271609
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-96G7"
1. 1097
/note="assembly_name:Contig14"
1198. 1253
/note="assembly_name:Contig15"
1354. 1743
/note="assembly_name:Contig16"
1844. 2940
/note="assembly_name:Contig28"
3041. 4209
/note="assembly_name:Contig42"
4310. 5533
/note="assembly_name:Contig43"
5694. 6894
/note="assembly_name:Contig45"
6995. 8292
/note="assembly_name:Contig49"
8393. 9864
/note="assembly_name:Contig52"
9985. 11951
/note="assembly_name:Contig53"
12052. 13101
/note="assembly_name:Contig56"
13202. 14843
/note="assembly_name:Contig59"
14944. 17128
/note="assembly_name:Contig61"
17229. 18762
/note="assembly_name:Contig62"
18863. 25092
/note="assembly_name:Contig63"
25193. 30277
/note="assembly_name:Contig64"
30378. 47187
/note="assembly_name:Contig65"
47288. 63647
/note="assembly_name:Contig66"
63748. 83396
/note="assembly_name:Contig67"
85497. 112995
/note="assembly_name:Contig68"
120096. 147799
/note="assembly_name:Contig69"
147900. 194953
/note="assembly_name:Contig70"
195054. 271609
/note="assembly_name:Contig71"

```

Query Match

```

Best Local Similarity 31.4%; Score 715.001; DB 1; Length 271609;
68.1%; Pred. No. 0.16;

```

```

Matches 1272; Conservative 0; Mismatches 490; Indels 106; Gaps 21;
QY 270 TTCCCTCTCTCTTCCAAACACTTCTATTTCTTGATTTCTATCTGGCTC-ATTTTAACT 328
DB 143362 TTTATTCAATCCTAGGAAGTCTTTCTTTTTTAAATTTCTGCTTGACACTTTTTTTCATT 143421
QY 329 CAGTAGTAGTTGTTTGGTTTCCATAAGTTTGTAAAGTTTCTGTTGTTCTGTTGTTGTT 388
DB 143422 GAGTAGAGAGTTGTTTCAGT-----GTCTACACTTCTGTTGTTTCTATT 143465
QY 389 GTTGTATCTAGATTTAAGCTGTGGTGCAGATAGGACATAGAGTATTATTTCATTTCT 448
DB 143466 AATGATATCCAGCTTTAATCCATGGTCTCATAGGGTCAGGGAGTTATTTCACATT 143525
QY 449 CTTTATCTGTGCGAGACTTGTGTTTGTGAAATATGATTCATTTTGGAG--AGTTTC 506
DB 143526 CTTGTATCTGTGAGACTTGTGTTTGTCTGAATATGATCATGATTTTGGAGACAATTC 143585
QY 507 ATAGGCTGTGACAAGAAGGTACAGTC-TTTGTGTTTGGTGAATAGTCTCTAAATATC 565
DB 143586 ATGAGGTGCTGAGAAGAGGTATATTCTTTGTGTTGGTGGAAATGTTTGTAGATATC 143645
QY 566 T-CTAGTCCACTTGGTTTATGACATCAGTATAGCTCCAGCAATTCCTGTTTTCGTTTTT 624
DB 143646 TGTGGGTTTCATCTGGTTTATAACATCAGTTAGCTCCAGCAATTTCTCTACTTAGTTTTT 143705
QY 625 GTT-----GAGATGACCTAACTGTTGGAGAAATGGGGTATTGAAGTAGCCCACTATCT 678
DB 143706 TTTTTTTCTGTTGCCCTGCTATTGGCAAGTAGGGTATTGAAGTCTCCCACTGCA 143765
QY 679 GTGTGTG-AGGTCAATATGTGATTTTAGCTGAGCTGTGCTTTGTTTATGAACCTGGGTG 737
DB 143766 CTGTGTGAAGTCAATGCATGATTTTAAGCTAACATAGTATCTGCTTATGAACCTGGGTA 143825
QY 738 ACATTGTGTTTGGTGCATAGACATTAAGAAATGCAATGTCCTCTTGGTGGAT-TTTCCTT 796
DB 143826 CTCTTGTGTTTGGGCATAGGTGTTAAGAAATCAAAACGCCATGTTGGTGGATCTTTCCT 143885
QY 797 TGATGCTATGTAGTATTCTTCCCAATCTCATCTGCTAGTTTGGTTTAAAGTCTA-TT 855
DB 143886 TGATGGCTCTGAAATACCTGCGCTCTCTCTCTGTTAGTTTGGTGTGATGGAATTT 143945
QY 856 AGTCAGATTAATAATGACTGTATCGGCTTGTCTTATAGGCCAATGCTGTAGATATCT 915
DB 143946 TGTCAACAGCAAAATGGCAACACCAAGATTGCTGTTAGGACTATCTCGTGAATACT 144005
QY 916 T-TTCCATCCTTTTACTCTAAGGTGATGTCATCCATGATAGTTGCTTTTTTGGATGC 974
DB 144006 TCTTGCAACCTTTATCTTAGGTAATGGCTACCTTGAATGTAAGTAAGTTTGTGGA 144065
QY 975 AGCAGTAGGATGGATCTTGTGTTTCATATCCATCTGTTACCCAGTATCTTTTCTAGAGA 1034
DB 144066 TGCAGCAGAATGAATCCTGTTTATCATCTTCTGTTAGACTGTATCTTTTATTTGGGT 144125
QY 1035 AATTAAGATCATTTAGTCAATGATTTGAGAATTAATCAATGACGAGTGTGTTGAGATCT 1094
DB 144126 AATTGAGA-----CCATTGATGTTGAGATATCAATGAGAATGGTTTGAATCT 144177
QY 1095 TGTATCTTCGCACTTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1154
DB 144178 TGTGATTTTGGTGGTGTGATCGGGGTGGGGGTGGGTGTGTGTGTGTGTGTGTGTGT 144237
QY 1155 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1214
DB 144238 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 144280
QY 1215 TCTTTTGAATTTTGGCTGGAATTTATTATTATTTCTTGAATGTGGTAAACAT 1274
DB 144281 TCTGGGATTTATTATTTCCAGTTTGTGTTTGTGTTTGTGTTTGTGTTGTTGTTGTT 144340
QY 1275 CTTTGAATGAAGTTTTCCTCTA---GCCCTTTTGTAGTCTGCAAT-TTGAAGATAGATA 1330
DB 144341 CTTTGAAGTGAATTAACCTTCCAAGTACTTCTGTAGGGCTCTATCTATTATATAAATA 144400

```

```
QY 1331 TTCCTTACATCGATTTTATCTTGAATGCTCTTCTTCTCCAACTATTGTGACAGAAAG 1390
Db 144401 TTGCTTACATTTGGCTTTATCATAAATGCTGTTTCTTCTCACTATTTGTGATTGAAG 144460
QY 1391 TTTTCTAAGTCAGTAGCTGGCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATC 1450
Db 144461 TTTTCTAGGTATAATAGTGTAGCTGGCATCTGTCTCTTTAAATCTGTAGCACAT 144520
QY 1451 TGTGAGGCGCTTCTTACATTTTCAGTTTCTATTGGAAGAGTCAGGTGTAAATCTCAATAC 1510
Db 144521 TGTCTAGGCGCTTCTTGTGTTTGTAGTCTC-----CATGAGAGTCTGTAG 144568
QY 1511 ATCTGCGTTTATATGTTAAATGCTTTTTCCTTTCCTTGCATCTTTTAATCTTTCTTTG 1570
Db 144569 GTCCTGCTTTATATGTTACTTGG-CTTTTTCCTTAAAGGCTTTTAATATCTTTTCTTTG 144627
QY 1571 TTCTATCTTTTGTAGTATGATTTATGATTTATGACATGCTGGGAGTCTCTTTCCGCTCCAAT 1630
Db 144628 -TAITGTGAGTCAGTGTGTTTGATTAATTAATGCTTAAGGGGCTTTCTTTCTGTCTCAAT 144686
QY 1631 CTATTTGTGTTTGTATGCTTCTTGTACCTTGTAGTCATCTTCTTCAAGTGTAGGA 1690
Db 144687 CTCCTTGTATTTCTAAATGC-----TACCTTTCTAGGCATCTCTTCTTTAGTGTAGGA 144740
QY 1691 AATTTTCTTTTGTGTTTCTTGAATAATTTTCCCTGCTTTTGTAGCTGCTTCTTCTTCCC 1750
Db 144741 AATTTTCTTCTATGATTTTGTGAATAATTTCTATGACTTTTGTAGCTGGTTCCTCCC 144800
QY 1751 C-----TTCCTCTATTTCTTGTGTTTGTGATAGTCTCTGCTTCTTCTGATG 1799
Db 144801 CCCCCCTCTATTTCTTATTAATCTTAGATTTAGTCTTTGCATCCCATATTTCTGATG 144860
QY 1800 TTTTATGCTGATTAATTTTATGACTTAACATTTCTTTTGAACCAAGGTATCAATTTCTTCT 1859
Db 144861 TTTTATCATCAGGGTCTCTTAGAGTTTACACTTTTCTTAACCAAGGTACCAATTTTCT 144920
QY 1860 ATCTGTCTTCTCAGTCCGTCAGATCTCTCTCTATCTCTGTTCTGTTCTGTCAGTGAGGCT 1919
Db 144921 ACTGTGCTTCTGATGCTCAACATTTCTCTTCCATCTTCTGCTGCTGCTGAG----- 144974
QY 1920 GTCTCTGAGTTCCTTGTGTTGTTCTTAATTTTTCATTTTCAGATTTTCTTCAAGTTTGG 1979
Db 144975 -----GTTTCTTTACCAAGTGTGTTAAATTTCTGTTTCAAGATTTTCTTCAAGTTTGG 145026
QY 1980 TTTTGTGTTTAAATCTATTTTCCATTTCCAGTCTGAGTCTGAGTGTGTTTACTCATTTCTCTCC 2039
Db 145027 TTTTGTGTTTAAATCTATTTCCATTTCCAGTCTGAGTCTGAGTGTGTTTACTCATTTCTCTCC 145085
QY 2040 CAGTATTTTACATTTTTCATGATTTTCTTTTAAATGATTTTATTCATTTCTTCTTCAAGGACCT 2099
Db 145086 CACTGTTTGTGTTTTCATGATTTTCTTTGTTAGATTTTACTCACTTTCTTCTTTTAAAGACCA 145145
QY 2100 TTTATGAA 2107
Db 145146 GTCAGGCA 145153

RESULT 20
LOCUS AC126868 247119 bp DNA linear HTG 10-MAY-2003
DEFINITION Rattus norvegicus clone CH230-50H7, WORKING DRAFT SEQUENCE.
ACCESSION AC126868
VERSION AC126868.4 GI:30521522
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 247119)
AUTHORS Muzny,D.Marie., Mettaker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
```

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregorogis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Loulseged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakolameh, O., Okwionu, G., Olariunsaogon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L., Puzio, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W. G., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Siason, I., Sitter, C. D., Smajda, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valdes, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

DIRECT SUBMISSION
Unpublished
2 (bases 1 to 247119)
Worley, K. C.
Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 247119)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:24819181.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

QY	1586	GATTGATTATTATGACCTGGGGAGTTCTTTTCGGTCCCAATCTATTGTTGTTTG	1645
DB	37765	GTATTGATTATATGTTGGAGGAGTCTTTCTTTTCTGGTCTCACTATTGTTGTTCTG	37706
QY	1646	TATGCTTCTTCTGACCTGATAGGACATCTTTTCTCAAGGTTAGGAAATTTTCTTTT	1705
DB	37705	TATGCTTCTTACAGCTTGACAGATCTCTCTTTTAAAG-TAGGGAATTTTCTTCTATA	37647
QY	1706	GTTCCTTGAAATATTTTCCCTGCTTTG-----ACCTGCCTTCTCCCTTCTCTA	1759
DB	37646	ATTCTGTGAAATATTTTCTGGGCTTTTGGGTTTCTTCTCTCTCTATCCCTTTGTCT	37587
QY	1760	TTCCCTTGGTTTGGCATAGTCTCTGCTCTCTGGAATTTTATGCCCTGGATTTATTT	1819
DB	37586	TAGATTGCTCTTTTATAGATTCAGATTTCTCTCCATGTTTGTGTTGAAATCTTTT	37527
QY	1820	AGACTTAACATTTCTTTTGACACAGGATCCATTTCTTCTATCTTGTCTTCACTGCTGA	1879
DB	37526	AGATTTAACATTTCTTTTGCCCATGATTAATTTCTTCTTATGTTAT-----CTGA	37476
QY	1880	GATTCTCTCTCTATCTCTGTTATCTGTCAGTCAGGCTTGTCTCTCAGGTTCTCTGTTG	1939
DB	37475	AATTATATTTTCCATCTCTTGATCTCTGTTGCGAAGCTTGCCTTTGTAGTTCTCTTTG	37416
QY	1940	GTTCTTAATTTTTCATTTCCAGATTTCTCTCAGTTTGGGTTTGTGTTTATTAATCTATT	1999
DB	37415	AAT-----TCCATAAATTTTCACTAAGTTTGTGTTGTTCTTTTATGTTCTTATT	37368
QY	2000	TCCACTTTTCCAGTCTGAAATGTTTCTACTCATTTTCTCTC-----	2039
DB	37367	TTCATTTTCCAGTCTTGAAATTTTACTCAATTTCTTCCACTTTTGTGTTGTTGCTTT	37308
QY	2040	-----CAGTATTATCATTTTTCATAGTTTCTTTTAAATGGAATTTATTTCTCT	2087
DB	37307	TTTATGATTTTCTCTCTCTCTGTTTCTCTCAATTTCTTTTAAAGAGATTTTAAATCTCT	37248
QY	2088	CTTCAAGACCTTTTATGAATCATAAATGATGTTAAAGTCTCTGCTTGTCTGCTTCAG	2147
DB	37247	CTTTAAGGCGCTCTATTACCTTCAT-GCAGGTGTTTAAAGGCTTTTCTTGTGCTTCAG	37189
QY	2148	CTATGTCATTTCTCAGGCGCTATTGTAATAGGT-----TTTACGAGGACATATT	2199
DB	37188	TTATGCTGGAAT-TGAGGACGTTGTTGTTAGGTGTTGCTGGGCTCTAATGTTGTCATAGT	37130
QY	2200	GTCTGTTGTTATGTTCTGTTTCTGTTTCTGTTTGGCATATAGACGCTGAGTTGGGATGA	2259
DB	37129	GCCCTGTTGTTATGATTTGTTT-TGGGCGTGGCATCTGTCGATCTGGAATTTGTTTGG	37071
QY	2260	TGTTATTTCTAGGCTGCTGAT 2279	
DB	37070	TTATAGGCTGGGCTGCTGAT 37051	
RESULT 21			
AC119553			
LOCUS			
DEFINITION	Rattus norvegicus clone CH230-11409, *** SEQUENCING IN PROGRESS		
ACCESSION	AC119553		
VERSION	AC119553.4		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
REFERENCE			
AUTHORS	Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anylebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,		

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleaveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, C., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunarane, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, J., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. B., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hui, Y., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Louised, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mathew, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakoelamen, O., Okwono, G., Olariunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivas, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sibson, I., Sitter, C. D., Smajd, D., Sneed, A., Sodergren, S., Song, X. Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wiecek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 259629)
Worley, K. C.
Direct Submission
Submitted (28-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 259629)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 23, 2002 this sequence version replaced gi:21903232.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWPR
Center clone name: CH230-11409

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 217607 bases at least Q40

Consensus quality: 220515 bases at least Q30

Consensus quality: 222444 bases at least Q20

Estimated insert size: 234666; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 3 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

* 1 5391: contig of 5391 bp in length

* 5392 5491: gap of unknown length

* 5492 39105: contig of 33614 bp in length

* 39106 39205: gap of unknown length

* 39206 259629: contig of 220424 bp in length.

FEATURES

source

1..259629

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-11409"

1..1145

/note="wgs_contig"

39206..40821

/note="wgs_contig"

Query Match

Best Local Similarity 31.4%; Score 714.803; DB 1; Length 259629;

Matches 1379; Conservative 0; Mismatches 537; Indels 144; Gaps 27;

Qy	277	TCCTCCAAACACTCTATTCTTGATTTCTATCTGGCTCAATTTTAACCTCAGTAGTG	336
Db	122307	TCCTCAATCAATCCATAAGGTTTTTAATTTCTGCTTGACCCACTTTTCATTCAGCAGAG	122366
Qy	337	AGTTGTTGGTTTCCATAAGTTTGAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTAT	396
Db	122367	A---GCTCAGCTTCTGTGAGTTTGAAGCTTCTGTTGTTCC---TGTGTTATCCGATCT	122420
Qy	397	CTAGATTAAAGCTGTGGGGTCCAGATAGACATAGAGTATTATTTCAATTGCTTTTATC	456
Db	122421	CTAATTAAATCCACGGTGTCTGATAGGATTCAGTACATATTCCAAAGTTTCTTGATC	122480
Qy	457	TGTCGAGACTTGCTTTGTTTGAATATGTATTCAATTTTGGAGAGTTTCATAGGGTGCT	516
Db	122481	TGTTGAGACTTG---TGTCFAAATTGGTGAATTTTCCAAAAGACCTATGACATGCA	122536
Qy	517	GACAAGAGGTACAGTCTTCTGTTTGGTGAATAGTCTGTAATATCT-CTAGGTCCA	575
Db	122537	GAAAAGAGGTATATCTTGTGTTGGTGAACGTTCTGTAATATCTGTAGTTCCA	122596
Qy	576	CTTGGTTATGACATCAGTTAGCTCCAGCAATTTCTGTTTCTGTTTGTTCAGATGAC	635
Db	122597	TTTGGTTGATGCTGCTCTCCAGCAATTTCTG-ITCAGCTTTTGTCTGATGAC	122655
Qy	636	CTAACTGTTGAGAGATGGGATATGAAGTAGCCCACTATCTGTGTGAGG-TCATA	694
Db	122656	CTATCTACTTGTGAGAAAAGAGATTGAAGCTTACTGCCATCCATATGTGAGGATCAATA	122715
Qy	695	TGTGATTTTACCTAGCTAGCTGCTGTTGTTTATGAACCTTGGGTGACATTTGTTGGTCA	754
Db	122716	TGTTATTAACTGAAGTGGTGTTCCTTTTACAACCTTGGGTGCCCTTATGTTGGGCA	122775
Qy	755	TAGACATTAAAGATTGCAATGCTCCTCTGGTGA-TTTTCTTGTGATGCCCTATGTAGTAT	813
Db	122776	TAGATGCCAAGATTGAATGTCATCTTGGTGAATTTTCTTTGATGAGTATGTAGTG	122835

Qy	814	TCCTCCAAATCTCACTGCTTAGTTTGGTTTTAAGTCT-ATTAGTCAGATATAAAATG	872
Db	122836	TCCTTCTATCTC-----TTGTTTGGTTGAAGTCTGTTTGTAGTAGTAAATG	122888
Qy	873	ACTGTATCGGCTTGCCTCTTAGGGCCATTGCTTAGAATATCTTTTCCA--TCCTTTTAC	930
Db	122889	GCTACACCACCTGGCTTCTTAGGGTCATTATATCTGAATATCTTTCCAATATCTTTAC	122948
Qy	931	TCTAAGGTGATGTCTATCC---ATGGTAGGTGCTCTTTTGGATGCAGCAGTAGGATGG	987
Db	122949	CCTGAGCAATGCTATCTCTTGAATTCAGGTATATTTATTCGATGCAGCAGAGATGG	123008
Qy	988	ATCTTGTTCATATCATCTCTTACCCAGATATCTTTTCTAGAAAATTAAGATCAT	1047
Db	123009	ATCTATTTCCACGGCCATCTTTAGTCTATTTGTTT-----AATTTGGGTAATA	123057
Qy	1048	GAGTCATTGATGTGAGAAATATCAATGACGAGTGTGTTGGATTCTTGTATCTTGCAC	1107
Db	123058	GAGACCATGGACTTGAGAGATACCAAGGACTAATGATTTGTTGATCTCTTACTTCTTC	123117
Qy	1108	TTGTGAAGTG	1167
Db	123118	TTCTTGTATTGTTGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	123177
Qy	1168	TGT	1227
Db	123178	TATGTAATGATGT-----TTCCCTTTTGAATTTGCT	123210
Qy	1228	GGCTCGAAATTTATTTATTTATTCATATTTCTTGAAATGGGGTAACATCTTTAGATTGAAG	1287
Db	123211	GTTCCTCAACTATTTCTTCTGTTTTCATGGGTATAGTCAACCTCTCTAGGTTGAAG	123270
Qy	1288	TTTTTCTCTAG--CCTCTTTAGTCTGCATTTGAGATAGATATCTTTTACATCTGAT	1345
Db	123271	TGTTTCATCTAGAACCTTCTGTAGGCTGGATTTGTAGACAGATATTTATTAATTTGCG	123330
Qy	1346	TTTATCTTAGAATGCTTTTCTTCTCAACTATTGTGACAGAAAGTTTCTTAAGTGCGAG	1405
Db	123331	TTTATCATAGATAAATC---TTCTCCATCTATGGTATGTAATGTTTC-CTGAGTAGAG	123386
Qy	1406	TAGTCTGGCTGACATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGCGAGGCTTCT	1465
Db	123387	TAGTCTCGTGGCACTGTGTTTCTTGGTATCTGATGAACATCTGCTCAGGCTCTCTCT	123446
Qy	1466	TACATTTTCAGTTTCTTATTTGAAAGTCAAGGTAACTTCTAATACATCTGCCCTTATAG	1525
Db	123447	GGCTATTAGATCCCACT-GAAGAGTCAATTTGTAATTTTAGTGTCTGCCCTTTATATG	123505
Qy	1526	TTAATTTGGTCTTTTTCCTTGCATCTTTTAAATATTTCTTTTCTTCTATATCTTTTGT	1585
Db	123506	TTGCTTGG--CCCTTTCCCTTGCACCTTTTAAATATTTCTTCTCTGTCGTATATTTAGT	123564
Qy	1586	GATTTGATTTATGCACTGTGGGAGTTCTTTTCCGGTCCAACTATTTTGGTGTGTTG	1645
Db	123565	GTTTGTATTATGTTGTTGAGGAGTCTTTTCTGCTCTCATCTATTTTGTGTCTG	123624
Qy	1646	TATGCTCTTGTACCTTGTATAGGCATCTCTTCTCAAGGTAGGAAATTTTCTTTTGTG	1705
Db	123625	TATGCTCTTGTACCTTGTACAGATATCTCTTTTAAAG-TAGGGAATTTTCTTCTATA	123683
Qy	1706	GTTTCTTGAAAATATTTTCCCTGCTTTG-----ACCTGCTTCTTCCCTTCTCTA	1759
Db	123684	ATTCTGTTGAAAATATTTTCTGGGCTTTTGGGTTTCTTCTCTCTATCTCTTCTTCT	123743
Qy	1760	TTCTTTGTTTGTATGATGCTCTGCTCTCTGGATGTTTATGCTGGATATTTT	1819
Db	123744	TAGATTTGCTCTTTTCAATAGGATTCAGATTTTCCCTCAGTTTGTGTTGAAATCTTT	123803
Qy	1820	AGACTTAACATTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCTACTGCCGA	1879
Db	123804	AGATTAACATTTCTTTGCCCATGTATTAATTTCTTCTATTTGAT-----CTGA	123854

QY 1880 GATTCTCTCTTCTATCTCTTCTATCTCTCTAGTCAGGCTGTCTCTCTGAGGTTCTCTGTGG 1939
 Db 123855 AATTATATTTCCATCTCTTGATCTCTGTGGCAAGCTTGCCTTCTAGTCTCTCTTG 123914
 QY 1940 GTTCTTAATTTTTCATTTCCAGATTTCCCTTCAGTTTGGGTTTGGTTTATATATCTATT 1999
 Db 123915 AAT-----TCCATAAATTTTCACTAAGTTTGGTGTCTTTATTTGGTCTTATT 123962
 QY 2000 TCCACTTTCAGGCTCGTAAGTGTCTTACTCATTTTCTCTCC----- 2039
 Db 123963 TTCAATTTTCAGGCTCTGAACAATTTTACTCATTTTCTCTCACTTTTGTGTGTCTGTT 124022
 QY 2040 -----CAGTATTTTACATTTTTCATAGGTTTCTTTAAGGATTTTATTCATTTCT 2087
 Db 124023 TTTATTGATTTTCTCCCTCTCTGTTTCTTCTGAATTTCTTTAAGAGATTTTATTAAGTTCT 124082
 QY 2088 CTTCAAGGACCTTTTATGAATTCATAAAATGATCTTAAAGTCTCTTCCCTTGGCTTCAG 2147
 Db 124083 CTTTAAAGGCGCTCTATTACCTTCAT-GCAGGTGTTTAAAGTCTTTTCTTGTGCTTCAG 124141
 QY 2148 CTATGTTTCATTTCTCAGGCGCTATTTGTAATAGGCT-----TTTAGCAGGACATATT 2199
 Db 124142 TTATGTTGAAT-TCAGGAGCTGTTGTGTAAGTGTCTGGGCTCTAATGTGTCATAGT 124200
 QY 2200 GTCTGTTGTTATTTCTGTTCTGTTTCTTCTTGGCATATAGACGCTAGTTTGGGATGA 2259
 Db 124201 GCGCTGTTGTTATTTGATTTGTTT-TGGGGCTGTCATCTGTCGATCTGGAATGTTGGTTGG 124259
 QY 2260 TTGTAATTTCTAGTCTGCTGAT 2279
 Db 124260 TTATAGGCTGCGGCTGCTGAT 124279

RESULT 22
 AC132786
 LOCUS Rattus norvegicus clone CH230-31M10, *** SEQUENCING IN PROGRESS
 DEFINITION Rattus norvegicus clone CH230-31M10, *** SEQUENCING IN PROGRESS
 AC132786
 AC132786.3 GI:25139109
 HTG: HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
 VERSION Rattus norvegicus (Norway rat)
 KEYWORDS Rattus norvegicus
 ORGANISM Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 235196)
 Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alebrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, J., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louisegh, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, N., Mahindartne, N., Mahmoud, M., Malloy, K., Mangum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwakweme, O., Okwunye, G., Olarnpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X. Z., Sorrelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabot, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 235196)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (04-SEP-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 235196)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:23693160.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GQXS

Center clone name: CH230-31M10

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 189575 bases at least Q40

Consensus quality: 193094 bases at least Q30

Consensus quality: 195346 bases at least Q20

Estimated insert size: 192812; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.


```

|||||
42260 TTGCTTCTATTCATTTCCATTTTCAGGCTTCGAATGGTTTATTTGTTTCATCCACTGTTGT 42319
QY TTTACATTTTCATAGGTTCTTTTAAATGATTTATTCATTTCTCTTCAAGGACCTTTTAT 2104
Db TTTCTATTTCTTGGCTTTCTTTTAAAGGATTTACTAATTT--TGCCCAATGTTGTTTA 42377
QY 2105 GAATTCATAAATGATGATTTAAGTCCCTGCTTGTCTTCAGCTATGTTGCAATTCAG 2164
Db 42378 TGTTTTCTCGAATTTCTTTGAATTCATTTGTTTATTCATTTCTCTATTATGTTTATAT 42437
QY 2165 GGCCTATTGTAATAGGTTTTCAGGAGGACATATGTCCTGG 2206
Db 42438 AGTTGGTTCTCTGCTGCTCAGCTATGTTGAGTATTCAGG 42479

RESULT 23
AC128004 247726 bp DNA linear HTG 20-NOV-2002
LOCUS Rattus norvegicus clone CH230-262L8, *** SEQUENCING IN PROGRESS
DEFINITION *** 3 unordered pieces.
ACCESSION AC128004
VERSION AC128004.3 GI:25139641
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 247726)
Murny D.Marie., Metzker M.Lee., Abramson S., Adams C., Alder J.,
Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,
Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed F.,
Biswalt K., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay J., Burch P., Burrell K., Calderon E.,
Cadenas V., Carter K., Cavazos I., Cesar H., Center A.,
Chacko J., Chavez R., Chen R., Chen Y., Chen Z., Chu J.,
Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
Davila M.B., Davis C., Davy-Carroil L., De Anda C., Dederich D.,
Delgado O., Denson S., Deramo C., Ding Y., Dinu H., Divya K.,
Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
Gebregorgis E., Geer K., Gill R., Grady M., Guerra W., Guvarra W.,
Gunaratne P., Haaland M., Hamill C., Hamilton C., Hamilton K.,
Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J.,
Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M.,
Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,
Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,
Karpachy S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,
Kowis C., Kraft C.I., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
Lorenshewa L., Lozsed H., Lozado R.J., Lu X., Ma J.,
Maheshwari M., Mahindartne M., Mahmoud M., Malloy K., Mangum A.,
Mangum B., Mapua P., Martin K., Martin R., Martinez E.,
Mawhney S., McLeod M.P., McNeill T.Z., Meenen E.,
Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S.,
Morgan M., Morris K., Morris S., Munidasa M., Murphy M., Nair L.,
Nankervis C., Neal D., Newton N., Nguyen N., Norris S., Parks K.,
Nwakoelameh O., Okwuonu G., Olarnpunsagoon A., Pal S., Parks K.,
Pasternak S., Paul H., Perez A., Perez L., Pfamkoch C.,
Plopper F., Poindexter A., Popovic D., Primus E., Pu L.L.,
Puzo M., Quiroz J., Rachlin E., Reeves K., Regier M., Reigh R.,
Reilly B., Reilly M., Reuter M., Richards S., Riggs F.,
Rivas C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J.,
Sanders W., Savery G., Scherer S., Scott G., Shatman S., Shen H.,
Shetty J., Shvartsbeyn A., Sisson I., Sitter C.D., Smajs D.,
Sneed A., Sodergren E., Song X.-Z., Sorelle R., Sosa J.,
Steimle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C.,
Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Usmani K.,
Valas R., Vera V., Villasana D., Waldron L., Walker B., Wang J.,

```

Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock G. and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 247726)
Worley, K.C.

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 247726)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 20, 2002 this sequence version replaced gi:23665065.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAJE
Center clone name: CH230-262L8
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 193410 bases at least Q40
Consensus quality: 198067 bases at least Q30
Consensus quality: 201676 bases at least Q20
Estimated insert size: 196651; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 245348: contig of 245348 bp in length
2 245349 245448: gap of unknown length
3 245449 246558: contig of 1110 bp in length
4 246559 246558: gap of unknown length
5 246559 247726: contig of 1068 bp in length.

Location/Qualifiers
1. 247726
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-262L8"
1. 1135
/note="wgs_end_extension
clone_end:T7"

FEATURES
source
misc_feature

[illegible]

RESULT 24
 AC095157/c
 LOCUS
 DEFINITION
 AC095157
 Rattus norvegicus clone CH230-8N16, WORKING DRAFT SEQUENCE, 3
 unordered pieces.
 AC095157
 HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa
 Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 245174)
 Muzny,D,Marie., Metzker,M, Lee., Abramzon,S., Adams,C., Alder,J.,
 Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
 Anyalebechi,V., Aoyagi,A., Aycof,M., Baca,E., Baden,H.,
 Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
 Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
 Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,B.,
 Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
 Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
 Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
 Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
 Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
 Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
 Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
 Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
 Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
 Gebregregis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
 Gunaratne,P., Healand,W., Hamil,C., Hamilton,C., Hamilton,K.,
 Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
 Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hughes,M.,
 Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
 Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
 Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
 Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
 Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
 Lorensuhwa,L., Loulseg,H., Lozado,R.J., Lu,X., Ma,J.,
 Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
 Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
 Mahoney,S., McLeod,M.P., McNeill,T.Z., Neenen,E.,
 Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
 Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
 Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
 Nwankwelenh,O., Okunolu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
 Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
 Plopper,F., Polindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
 Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Sosa,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wleczyk,R., Woodden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zhao,S., Dunn,D., von
 Niederhauser,A., Weis,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstein,G. and Gibbs,R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 245174)
 Morley,X.C.
 Direct Submission
 Submitted (16-Sep-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 245174)

AUTHORS
 TITLE
 JOURNAL
 COMMENT

Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On May 9, 2003 this sequence version replaced gi:24817806.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GCUR
 Center clone name: CH230-8N16
 ----- Summary Statistics
 Assembly program: Atlas;
 Consensus quality: 225865 bases at least Q40
 Consensus quality: 227968 bases at least Q30
 Consensus quality: 229513 bases at least Q20
 Estimated insert size: 237268; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 241598: contig of 241598 bp in length
 * 241599 241698: gap of unknown length
 * 241699 242839: contig of 1141 bp in length
 * 242840 242939: gap of unknown length
 * 242940 245174: contig of 2235 bp in length.

FEATURES
 source
 1..245174
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-8N16"
 2104..3008
 /note="clone boundary
 clone_end:T7
 site:EcoRI
 end sequence:BH347120"
 2142..2962
 /note="clone boundary
 clone_end:T7
 site:EcoRI
 end sequence:BH347120"
 239787..241598
 /note="wgs end_extension
 clone_end:T7"

misc_feature
 2104..3008
 /note="clone boundary
 clone_end:T7
 site:EcoRI
 end sequence:BH347120"
 239787..241598
 /note="wgs end_extension
 clone_end:T7"

misc_feature
 2142..2962
 /note="clone boundary
 clone_end:T7
 site:EcoRI
 end sequence:BH347120"
 239787..241598
 /note="wgs end_extension
 clone_end:T7"

misc_feature
 2104..3008
 /note="clone boundary
 clone_end:T7
 site:EcoRI
 end sequence:BH347120"
 239787..241598
 /note="wgs end_extension
 clone_end:T7"

Query Match 31.2%; Score 710.102; DB 1; Length 245174;
 Best Local Similarity 69.3%; Pred. No. 0.15;
 Matches 1307; Conservative 0; Mismatches 479; Indels 99; Gaps 26;

435 ATTATTTCAATTGCTTTTATCTGTCGAGACTGCTTTGTTTGAATATGTTCAATT 494
63549 AATTTTCAATCTCTTTGTAICTCTAAGACTTCTGTTTGCTGTAATATGTAATCAATT 63490
495 TTGAG--AGTTTCATAGGCTGTCACAGAAGGTACAGTC-TTGTGTTTGTGGTGAAT 551
63489 TTGAGAGAAGTTCCATGAGATGCTAAGAAGAAGGTATATTCTTTTGTGTTGGTGAAT 63430
552 AGTCGTAAATATCTCTAGGTCACCT---TGTTTATGATCATGATGATGCTCCAGCAATTT 608
63429 GTTCTGTACATATCTGTTGGGTTCACTGGGTATATAGCAATGTTAGCTGTAGCAATTT 63370
609 CTCTGTTTCTGTTTCTGATGAGCACTAACTCTGGAGAGAAATGGGGTATTGAAGTAG 668
63369 CTCT-----AGTTTATCTAGATGCCCTGCTATTAGCAAGAGTAGGGTATTGAAGTCT 63315
669 CCCACTATCTGTGTG-AGTCAATATGTAATTTAGCTGTAGCTGTGCTGTTTATG 727
63314 CCCACTGTCAAGATGTGAAGTCAATGTGTAATTAAGCTAAAGTAGTGTATGCTTTATG 63255
728 AACTTGGGTGACATTGTTTGGTGGCATAGACATTAAGAATTGCAATGCTCTCTGGTGG 787
63254 AACTTCAGTACCTTGTGTTGGGTCATAGATGTTAAATCAAAATGTCATCTGGTGG 63195
788 A-TTTTCTTTGAGCTGATGATATCTTCCAACTCATCTGCTAGCTGTTTGGGTTT 846
63194 ATTTTCTCTTGACAAGTATAAAGTATTCTTCCCTGCTCTCTGTTGTTAGTTTGGTTG 63135
847 AAGCTATTAGTCAGATTAATAATGACTGTATCGGCTTCTCTTCTAGGGCCATTGCTT 906
63134 ATGTAT-TTTGTGACAAACAAAATGGTATACAGCTTACTTATTAATACCCATTGCTT 63076
907 AGAATACTTTTCC---ATCCTTTTACTCTAAGGTGATGCTATCCATGCTAGGTTGTCT 963
63075 AGAGTACCTTCTCCAAACACTTTATACCTTACATAATGGCTATCCTTGATGTTGAAGTG 63016
964 TTTTGGATGAGCAGTAGATGATCTGTTGTTTCAATCCATCTGTTACCCAGTATCT 1023
63015 TATTCTTGGATGAGCAGGAGTGAATCCTGTTTATACATCCATTCTGTTAACTGTTAT 62956
1024 TTTTCTAGAGAAATTAAGATCAATTGAGTCATTGATGTGAGAATTATCAATGAGCAGTGT 1083
62955 TTTAATTGGTCAATTGAGA-----CCACTGATGTTGGAGATATCAATAGAATGT 62904
1084 TTGAGATCTGTATATCTGCACTTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1143
62903 TTTCTGATCTTGTGATTGTTGGTGTG-GTAGTGATGTTGGTGTGTGTGTGTGTGTGTGTGT 62845
1144 TGT 1203
62844 GGT 62788
1204 GTTCTCTCCCTCTTTGATTTTGGCTGGAATATTATTATT---ATTCAATTTTCTTGA 1261
62787 GTT-----TTCCGTGCTGAGATTAATTATTCCCTTTTTCACCTTTG 62744
1262 ATGTTGGGTAAACATCTTTAGATGAAGTTTCTCTCTA---GCCTTCTTTAGGCTGCAAT 1318
62743 GTGTGGGAGGCTCTTTAGACAGAGTTAACTTATAGTAATTCAGTAGAGCTTGAAT 62684
1319 TGAAGATAGATATTCTTTTACATCTGATTTTATCTTTAGAACTGTTCTTTCTCCAACTAT 1378
62683 TATATATAAATATTGCTTGCAATTAGTTTATCATATAAAGTCTGTGTTCTTCTCATCTAT 62624
1379 TGTGACAGAAAGTTTTTCTAAGTCAGTAGTCTGCGCTGACATCTGTAGTCTCTTGGAGT 1438
62623 TGTGATTGAAAG---TGCTGGGTATAACAATGACAGGCTGGCATCTGTGGTCTCTTAGCCT 62567
1439 CTGTAGCACATCTGGAGGCTCTTACATTTTGTAGTTTCTATTCGAAAGTCAGGTG 1498
62566 CTGTAGCACATTTGGCTAGGCTCTTTTGTCTTTTAGAGTCTTCCA-TGAAAGTCAGGTG 62508
1499 TAAATCTAATACATCTGCTTTTATATGTTAATGTTCTTTTCCCTTGTGATCTTTTAAAT 1558

62507 TAGTCTGATAGTCTGCTTTGTAATGTAATGTTTCTCCCTGACGCTTTTAAAT 62448
1559 ATTCTTTC-TTGTGTCATACATTTTGTAGTGAATTTGATGATGCTAGTGGGAGTTTCT 1617
62447 ATTCTTCTTTTCTGTCATACATTCAGTGTGTTTGTATTTATGTCACAGGAGACTTCT 62388
1618 TTTCCGGTCCAATCTATTTTGGTGTGTTTGTATGCTTCTTGTACCTTGTATAGGCACTCTCTT 1677
62387 TTTCTGGCTCAATCTATTTGGTATTCTGTATGCT-----TACCTTTCTAGGCACTCTCTT 62334
1678 CTCAAGGTTAGAAATTTTCTTTTGGTGTGTTTCTGAAATATTTTCCCTGCTTTTGAC 1737
62333 CTATTATGTAGAAATTTTCTTCTATGATTTTGTGAAATATTTTCTGTGCTTGTGAC 62274
1738 CTGCTCTCTTCC-CTTCTCTCTATTCC-----TTTGGTGTGTCATAGTGTCT 1783
62273 CTGGGATCCCCCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 62214
1784 TCTGGCTTCTCGATGTTTATGCTGGAATTTTATAGACTTTAAGCAATTTCTTTGACCAA 1843
62213 CCATAATTTCTGGACATTTTATACAGGAGTTTCTTAGAGTAAACATTTCTTTAAACAA 62154
1844 GGTATCCATTTCTTCTATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1903
62153 GGCATTCATTTCTTCTTCTTCT 62096
1904 TCTGTCAGTGAGGCTGTCTCTGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1963
62095 TCTGT-----TGCCCTGAGTCTTCTTCCGGTGTCTTAAATTTCTCAATTTCCAGA 62045
1964 TTTCTTCAAGTTGGGTTTGTATTAATCTTATTTCCACTTCCAGTTCCTGAGTCTGAAATGTT 2023
62044 TTTCTCTCAAGTTGGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 61985
2024 TTTACTCATTTCTCTCCAGTATTATTAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2083
61984 TTTATCTA-TTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 61927
2084 TCTCTTCTCAAGGACCTTTTATGAATTCATTAATATGATGTTAAGTCTTCTGCTTGTGCT 2143
61926 TCTCTTTTAAAGATCTCTATCTATCTATGTCATTAAGGCTGTTTGGGCTTATTTTGTGTT 61867
2144 TCACTATGTTGATCTCTCAGGCT-----TATGTAATAGGTTTACAGGAGAC 2194
61866 TCACTATGTTGATCTGAAATATTCAAGGCTGCTGTTGTTAGGTTGTTGTTGTTGTTGTTGTT 61807
2195 ATATTCTCTGTTGTTTATTGTTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2254
61806 ATATAGTCTGATCTGTTTATGTTTGTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 61749
2255 GATGATTGAATCTTAGGTTGCTGAT 2279
61748 AATGATTGAATTTTCAAGTGTTCAT 61724

RESULT 25

AC113514
LOCUS 167223 bp DNA linear HTG 21-OCT-2002
DEFINITION Mus musculus clone RP23-475D11, WORKING DRAFT SEQUENCE, 9 ordered
pieces.
ACCESSION AC113514
VERSION AC113514.4 GI:24182231
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 167223)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Birren, B., Nussbaum, C. and Lander, E.
JOURNAL Mus musculus, clone RP23-475D11
REFERENCE 2 (bases 1 to 167223)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Canarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooper,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gadyne,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Mathews,C., McCarthy,M.,
Mcwan,P., McKernan,K., Meldrum,J., Meneus,L., Minova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembeck,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (01-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 167223)

AUTHORS

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Canarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gadyne,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J.,
Mathews,C., McCarthy,M., Meldrum,J., Meneus,L., Minova,T.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembeck,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (21-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Oct 21, 2002 this sequence version replaced gi:21326539.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 10572: contig of 10572 bp in length
* 10573: gap of 100 bp
* 10673 12088: contig of 1416 bp in length
* 12089 12188: gap of 100 bp
* 12189 13362: contig of 1174 bp in length
* 13363 13462: gap of 100 bp
* 13463 12741: contig of 9279 bp in length
* 12742 22841: gap of 100 bp
* 22842 27991: contig of 5150 bp in length
* 27992 28091: gap of 100 bp
* 28092 40867: contig of 12776 bp in length
* 40868 52499: contig of 11532 bp in length
* 52500 52599: gap of 100 bp
* 52600 154378: contig of 101779 bp in length
* 154379 167223: gap of 100 bp
* 154479 167223: contig of 12745 bp in length.

FEATURES

source

1..167223
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-475D11"
/clone_lib="RPCI-23 Female Mouse BAC"

misc_feature

1..10572
/note="assembly_fragment
clone end:SP6
vector side:left"

misc_feature

10673..12088
/note="assembly_fragment"
12189..13362
/note="assembly_fragment"

misc_feature

13463..22741
/note="assembly_fragment"
22842..27991
/note="assembly_fragment"

misc_feature

28092..40867
/note="assembly_fragment"
40968..52499
/note="assembly_fragment"

misc_feature

52600..154378
/note="assembly_fragment"
154479..167223
/note="assembly_fragment"

misc_feature

clone end:17
vector side:right"

Query Match 31.0%; Score 707.202; DB 1; Length 167223;
Best Local Similarity 67.5%; Pred. No. 0.12;
Matches 1385; Conservative 0; Mismatches 538; Indels 130; Gaps 29;

QY	275	TCCTCTCCAAACACTTCTATTTCTGATTCATCTTGGCTCATTTTAACTCAGTAG	334
Db	71989	TGAGTTTAGAAGAGTCTTTAAATTTCTTTTCTGTTTGCACATTTT--TTTCTGTTTGA	72046
QY	335	TGAGTTTGTGTTT-GGTTTCCATAAGTTTGTAAAGTTTCTGTTTCTGTTTGTGTTGTGT	393
Db	72047	ACAGTTGTTTGTAGTTTTCATGATTTTGTAAAGCTTCTATTATTTTCTGTTTGTGTTTGA	72103
QY	394	TATCTAGATTTTAAAGCTTGGTGGTCAGATAGACATAGATATTTCAATTGCTTTT	453
Db	72104	TATCCATCTTTAATCCGTGGTGGCTGAGAAATAGATAGTGTAGTTCAA-TGTTCTTGT	72162
QY	454	ATCTGTCCAGACTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT	512
Db	72163	ATCTGTGAGACTTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT	72222

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. Gaps between the contigs

```
QY 513 TGCTGACAAGAGTACAGT---CTTTGTGTTTTGGTGAATAAGTCTGTAAATATCT-CT 568
Db 72223 TATAGAGAAGAACTGTATATTTCTTTTGTGTTGGGTGAGATGTTCTGTAAATATATATT 72282
QY 569 AGGTCCACTTGGTTTATGACATCAGTTAGCTCCAGCATTTCTCTGTTTCTGTTTGTG 628
Db 72283 AGGTCCACTTGGTTTATGACATCAGTTAGCTCCAGCATTTCTCTGTTTCTGTTTGTG 72341
QY 629 AGATGACCTAACTGTTGGAGAGAAATGGGGTATTAAGTAGCCACATATCTGTGTGAGG 688
Db 72342 AGATGACCTGTTGTTATGTTAGTGTGTAATGAAGTCTCCCAACCGTCACTGTGTGAGG 72401
QY 689 TCAATATGCAATTTAGCTGTAGCTGTCTGTTTATGAAGCTTGGTGACATTTGTGTTT 748
Db 72402 GTA-----AGTAGTGTATTTTACAAGGTGGGTG-----72432
QY 749 GGTCATAGACATTAAGAAATGCAATGCTCTCTCTGTGGTGA-TTTTCCCTTTGATGCCATG 807
Db 72433 ---GCATAGATGTTAATAAGAAATATCATCTTGGTGAATTTTCCCTTTGATGAGTAG 72489
QY 808 TAGTATTTCCCAATCTCATCTGCTTAGTTTGGTGTAAAGTCTAT-----TAGT 858
Db 72490 TTGTGCTCTTCCCTGCTCTCTAATAATAGTTTGGTCTGAAGTCTGTTTATTTGATATC 72549
QY 859 CAGATATTAATACTGATCTGCTTCCCTTCTTAGGCCATTTGCTTAGAATATC----914
Db 72550 AAAATAGCTACATACTATAGGACCTTTTCTTAGGTCCATTTGTTTGAATAATCTTT 72609
QY 915 -TTTTCCATCTTTACTCTAAGTGATGTCATC---CATGGPAGTGTCTCTTTTGG 970
Db 72610 TTTTGTGAACCTTTACCTTAGGTAAATTTCTATCTTTGATTTGAAGTATATTTCTGG 72669
QY 971 ATGCAGAGTAGTAGATCTGTTTTCATATCTGTTACCACTATCTTTTCTA 1030
Db 72670 ATGCAGCAAAAGGATGATCTGTTTGGCACTCTTTCTATTAGACTGTGCTTTTATT 72729
QY 1031 GAGAAATTAAGATCATTTAGTCAATGTTTGAAGATTAATCAATGAGCAGTGTTTGGGA 1090
Db 72730 GGGGATTTAGACCATTA-CCATTTAGATTTAGAGAGATCAGTGCCCAATTAATGTTGA 72788
QY 1091 TTTCTGTTATCTTGCACTGTGAAGTGTGTGTG-TGTGTGTGTGTGTGTGTGTGTG 1149
Db 72789 TTTCTGTTGTTTGTGTTTATTTATTTGTTGGTGTGTGTGTGTGTGTGTGTGTGTGTG 72848
QY 1150 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1209
Db 72849 TTTATGTTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 72905
QY 1210 TCCCCTCTTTGATTTTGGCTGGAATTAATTTATTTATTTATTTATTTATTTATTTAT 1269
Db 72906 -----TTAGATTTGTTTATTTTCCATGTTTGTGTTTATAGTT 72943
QY 1270 AACATCTTAGATTAAGATTTTCTCTAG--CCTTCTTTAGGCTGCAATTTGAGATAG 1327
Db 72944 TTTCTCTTAGATTTGAGTTTCTCTCCAGACCTCTGCAGGACCTGATTTGATAGATAG 73003
QY 1328 ATATCTTTATCATCTGAT-TTATCTTGAATGCTTTCTTTCTCCCAATTTGTCAGAC 1386
Db 73004 ATCAATATTTGTTTAAAGTGTTTTAAAGAAATATCTTATTTCTTCCATCTGACTG 73063
QY 1387 AAAGTTTCTTAAGTGCAGTAGTGTGCTGACATCTGATCTCTTGGAGTCTGTAGCA 1446
Db 73064 AGAGTCTTGTCTGGGTATAGTAGTGTGGCTGACATCTGTGTTTCTTATAGTCTGTCAGCA 73123
QY 1447 CATCTGTCAGGGCTCTTCAATTTTGAAGTTTCTATTTGAAAGTCAAGGTGTAATCTA 1506
Db 73124 CATCTGCCAGATCTTCTGCTTTTAGT-----CCATTTAGAGTCAAGGTGTAATCTC 73180
QY 1507 ATACATCTGCCCTTATATGTTAAATGGTCTTTTCTTCCCTGCACTTTTAAATCTTTC 1566
Db 73181 ATAGTCTGCTCTTTGATGTTACTGTGTC-TTTTCCCTTGTAGCTTTTAAATATCTTCC 73239
```

```
QY 1567 TTTGTTCTATACITTTAGTGAITGATTTATATGCACTGTGGGAGTTTCTTTCCGTC 1626
Db 73240 TTTGTTCTGTACATTAATATTTGATTTAAGTGTGAGGGGCTTTCTTTTCTGTT 73299
QY 1627 CAATCTATTTGTTGTTGATGCTTCTGTACCTTTGATAGGCATCTCTTTCTCAAGTT 1686
Db 73300 CAATCTATTTAGTGTCTATATGCTTCTGTACCTTTATAGCACTCATCTTTAGTT 73359
QY 1687 AGGAAATTTTCTTTTGGTTTCTTGAATAATATTTCCCTGCTTTTGGACCTG-----1740
Db 73360 AGGAAATTTTCTCTCATATTTTGAATAATGTTTCTGGGCTTTTGGAGCTGAAATTC 73419
QY 1741 -CCTTCTTCCCTCTCTCTATTTCC-----TTTGGTTTTGCAATGCTCTCT 1786
Db 73420 TCTTCTCTCTCTTCTCTCTCTCTCTATTTTCTTATGATTTGCTTTTCACTGGGCTCA 73479
QY 1787 GGTCTCTGATGTTTATGCTGCTGATTTATTTAGACTTAACATTTCTTTGACCAAGT 1846
Db 73480 AATTTCTGTTGTTTGTGTC---AGCTTTTATGATTTAATATTTCTTTTGA---GGTGT 73534
QY 1847 ATCCATTTCTCTATCTGCTCTTCACTGCTGAGATTTCTCTCTCTATCTCTTATTTCT 1906
Db 73535 ATTTATTTCTCTATGTTTCTTATTTATGATTTCTGTTTCCATTTTCAAGTATTTGAACAGT 73714
QY 1907 GTCACTGAGGCTTGTCTCTGAGGTTCTCTGTTGGTCT---TAAATTTTCAATTTCCAGA 1963
Db 73595 ATAGTGAAGCTTGCCTGTTAGTTACGTTTCAATTTCTAAATTTCTTTTCAATTTTCA 73654
QY 1964 TTTCTTCAAGTTGGTGTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2023
Db 73655 ATTCCTCAGTTGGGTTTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 73714
QY 2024 TTACTCATTTTCTCTCCAGTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2083
Db 73715 TTATTTCA-TTTTCTTCCACCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 73773
QY 2084 TCTCTTCAAGCACTTTTATGAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2143
Db 73774 TCTCTTTAAGGATCTCTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 73833
QY 2144 TCAGCTATTTTGA---TTCTCAGGCTTATTTGTAATAGGTTTTAGCAGGACATATTTG 2200
Db 73834 TCAGCTATTTTAAAGGCTCTGTTGTTAGGATGAGGCTTCCATGAGACAAATTTG 73893
QY 2201 TCTGTTGTTATTTGTTCTGTTTGTCTTTTGGCATATAGAGGCTGAGTTTGGGATGAT 2260
Db 73894 CCCCAG-TGTTATTAATGTTGTTTTATGTTAGCATCTAGGTATCTGGAGTTGGGATGAT 73952
QY 2261 TGTAACTTAGGT 2273
Db 73953 TGTAGGCTAGTT 73965
```

RESULT 26

```
AC110100 250188 bp DNA linear HTG 20-NOV-2002
LOCUS Rattus norvegicus clone CH230-212E13, WORKING DRAFT SEQUENCE, 2
DEFINITION unordered pieces.
AC110100
VERSION AC110100.6 GI:25139572
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Chordata; Vertebrata; Euteleostomi;
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Rattus.
1 (bases 1 to 250188)
AUTHORS Kuzny, D., Marie, M., Metzger, M., Lee, S., Adams, C., Alger, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebech, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
```

Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, Y., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gregoire, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Guaratone, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwa, L., Loulseged, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwuonu, G., Olarunpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, I., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Kelly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rokey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steime, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished
2 (bases 1 to 250188)
Worley, K.C.
Direct Submission
Submitted (10-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 250188)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (26-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 20, 2002 this sequence version replaced gi:23270205.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly ('a' 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCY

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GOFP
Center clone name: CH230-212E13
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 227481 bases at least Q40
Consensus quality: 230202 bases at least Q30
Consensus quality: 232076 bases at least Q20
Estimated insert size: 239268; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 33011: contig of 33011 bp in length
* 33012 33111: gap of unknown length
* 33112 250188: contig of 217077 bp in length.
* Location/Qualifiers
* 1..250188
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-212E13"
* misc_feature 31752..33011
* /note="wgs contig"
* misc_feature 33112..35207
* /note="wgs contig"
*
Query Match 31.0%; Score 706.902; DB 1; Length 250188;
Best Local Similarity 69.2%; Pred. No. 0.16;
Matches 1305; Conservative 0; Mismatches 481; Indels 99; Gaps 26;
*
QY 435 ATTATTCAATGCTCTTTTATCTGCGAGACTGCTTTGTTGTAATATGATTCATT 494
Db 48760 AATTTTCAATCTTCTCTGTAAGACTTGTCTGCTGAATATGATTCATT 48819
*
QY 495 TTGGAG--AGTTTCATAGGGTCTCACAAGAGGTACAGTC--TTTGTCTTTTGGTGAAT 551
Db 48820 TTGGAGAAAGTTCCATGAGATGCTAAGAGAGAGGTATATTCTTTGTTGTTGGAT 48879
*
QY 552 AGTCTGTAATATCTCTAGTCCACT---TGTTTATGACATCATGTTAGTCCAGCAATT 608
Db 48880 GTTCTGTACATATCTGTTGGTTTCATCTGGTGTATATAGCATTTGTTAGCTAGCAATT 48939
*
QY 609 CTTCTGTTCTGTTTGTGATGATCACTTGTGGAGAGATGGGGTATTGAAGTAG 668
Db 48940 CTCT-----AGTTTATCTAGATGCCCTGCTATTAGCAAGTAGGTATGAAGTCT 48994
*
QY 669 CCCACTATCTGTGTG-AGGTCAATATGATTTTATGCTAGCTGCTGTTGTTATG 727
Db 48995 CCCACTGTCAGGATGGAAGTCAATGTGATTTAAGCTAAGTAGTGTATGCTTTATG 49054
*
QY 728 AACTTGGGTGATGCTGTTGGTCATAGACATTAAGATTCGAATGCAATGCTCTTGGTG 787
Db 49055 AACTTCAAGTACCTTGTGTTGGGTCATAGATTTAAAAATCAAAATGTCATCTTGGTG 49114
*
QY 788 A-TTTTCTTGTATGCCCTATGATGATTTCTTCCCAATCTCATCTGCTTAGTTTGGGTTT 846
Db 49115 ATTTTCTTGTGACAAAGTATAAGTATTTCTCCCTGCTCTCTCTGTTAGTTTGGTTG 49174
*
QY 847 AGTCTATAGTCAGATATTAAATGACTGATCGGCTTGTCTTCTAGGGCCATTTGCTT 906
Db 49175 ATGTAT-TTTGTGACAAACAAATGGCTATTACAGCTTACTTATTATACCCATTTGCTT 49233

QY 907 AGAATATCTTTTC---ATCCTTTTACTCTAAGGTGATGCTATCCAGTGTAGTGTCT 963
Db 49234 AGAGTACCTTCTCCAAACACTTTATACCCCTACATAATGCTATCCTTGATGTAAGTG 49293
QY 964 TTTTGTGATGACAGTAGATGATCTGTTTTCATATCCATCTGTTACCCAGTATCT 1023
Db 49294 TATTTCTTGGATGACAGGATGATCTGTTTTCATCCATCTGTTTAACTGTATAT 49353
QY 1024 TTTTCTAGAGAAATTAAGATCATTTGAGTCAATTTGATGTTGAGAAATTAATCAATGACGATGT 1083
Db 49354 TTTAATTTGGTCAATTGAG-----CCACTGATGTTGGAGATATCAATAAGAAATGT 49405
QY 1084 TTGTGGATCTTGTATCTTGTCACTTGTGAAGTGTGTTGTGTGTGTGTGTGTGTGTG 1143
Db 49406 TTTCTGTATCTTGTGATTTTGGTGGTG-GTAGTGAAGGTGGTGGTGGTGGTGGTGGT 49464
QY 1144 TGT 1203
Db 49465 GGTGGTGGT 49521
QY 1204 GTTCTCTCCCTCTTTTGAATTTTGGCTGGAATTAATTAAT--ATTCATATTTTCTGA 1261
Db 49522 GTT-----TTCTGGTCTGAGATTAATTAATTCCTTTTTCACCTTTG 49565
QY 1262 ATGTGGGTAAACATCTTTAGATTTGAAGTTTTCCTTA---GCCTTCTTTAGGTCTGCAT 1318
Db 49566 GTGTGGGAGCCTCTTTAGACGAGATTAACCTTATAAGTAATTTAGTAGAGCTGAT 49625
QY 1319 TGAAGATAGATATCTTTACATCTGATTTATCTTGAATGTCCTTTCTTCTCAACTAT 1378
Db 49626 TATATATAATATGCTTGCATTTAGTTTATATCAATAAATGTCCTTTCTTCTCATCTAT 49685
QY 1379 TGTGACAGAAATTTTCTAAGTGCAGTAGTCTGGCTGACATCTGAGTCTCTTGGAGT 1438
Db 49686 TGTGATTTGAAAG---TGTGGGTATTAACAATGCAAGGCTGCATCTGGTCTCTTAGCCT 49742
QY 1439 CTGTAGACATCTGTGAGGCGCTTCTTACATTTTGTAGTTTCTATTTGAAAAGTCAGGTG 1498
Db 49743 CTGTAGACATCTGTGAGGCTTCTTCTTGTCTTTTAGAGTCTCA-TGAAAAGTCAGCTG 49801
QY 1499 TAATCTAATAACATCTGCTTTATATGTTAATGCTTTTCTTCTTCTTCTGATCTTTTAA 1558
Db 49802 TAGTCTGATAGTCTGCTTTGTATGTTACTTCTTTTCTTCTTCTGAGCTTTTAA 49861
QY 1559 ATTCTTCTC-TTTGTCTATCTTTTAGTGAATTCATTTATATGACATGCTGGAGTTTCT 1617
Db 49862 ATTCTTCTTCTTCTTCTGATCTAGTCTGTTTGTATGTTTGTATGTTTGTATGTTTGT 49921
QY 1618 TTTCCGGTCCAACTATTTGTGTTTGTATGCTTCTTGTACCTTGTATGCGATCTCTTT 1677
Db 49922 TTTCTGGCTCAATCTATTTGTATTTCTGTATGCT-----TACCTTTCTAGGATCTCCT 49975
QY 1678 CTCAGGTTAGAAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTGAC 1737
Db 49976 CTTTATGTAGAAATTTTCTTCTATGTTTCTGAAATATTTTCTGTCGCTTTGAC 50035
QY 1738 CTGCTCTCTTCC-CCTTCTCTATTC-----TTTGGTTTGTGATGTC 1783
Db 50036 CTGGATCCCTCTTCTCTCTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 50095
QY 1784 TCTGGCTCTCTGATGTTTATGCTGATTAATTTTATGATTAATCTTCTTCTTCTTCTG 1843
Db 50096 CCATATTTCTCTGGACATTTTATACAGGATTTCTTATAGATTAATCTTCTTCTTCTT 50155
QY 1844 GGTATCAATTTCTTCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTGAT 1903
Db 50156 GCAATCTTCTTCTATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 50213
QY 1904 TCTGTAGTGGCTTCTCTGAGGTCTCTGTTGGTCTTCTTCTTCTTCTTCTTCTTCTG 1963
Db 50214 TCTGT-----TGCCCTGGAGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 50264
QY 1964 TTTCTTCTAGTTTGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2023

Db 50265 TTTCCTCAGTTGGTTCCTTTTATTAGTTTATTCTTACATTCAGATCTTGAAGTCT 50324
QY 2024 TTACTCATTTTCTCCAGTATTTTACATTTTTCATAGTTTCTTAAATGATTTATTTCT 2083
Db 50325 TTATTCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 50382
QY 2084 TCTCTTCAAGACCTTTTATGAATTCATAAATGATGTTAAGTCTTCTTCTTCTTCTT 2143
Db 50383 TCTCTTTTAAAGATCTTCTATCATTTGTATAAGGCTGTTTGTAGGTTCTTATTTTGT 50442
QY 2144 TCAGTATGTTGCAATTCAGGCTC-----TATTGTAATAGGTTTATAGCAGGAC 2194
Db 50443 TCAGTATGTTGCAATTCAGGCTCCTGTTGTAGGTTGTTGTTGTTGTTGTTGTTGTTG 50502
QY 2195 ATATTGCTGCTGTTTATTTGTTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2254
Db 50503 ATATAGTCTTGAATTTTATTTAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2254
QY 2255 GATGATGTAATTTCTAGGTTCTGAT 2279
Db 50561 AATGATGTAATTTCTAGGTTCTAT 50585
RESULT 27
AC129113 168169 bp DNA linear HTG 19-NOV-2002
LOCUS Rattus norvegicus clone CH230-34207, WORKING DRAFT SEQUENCE.
AC129113
AC129113.4 GI:25072866
HTG: HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 168169)
AUTHORS Muzny, D. Marie., Metzker, M. Lee., Abramzon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Y., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Delgado, O., Denson, S., Denson, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karnathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, X., London, P., Longacre, S., Lopez, J.,
Lorensu, L., Louised, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Kwaokelameh, O., Okwuon, G., Olarnpunsagoon, A., Pal, S., Pankoch, C.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pflankoch, C.,
Plopper, F., Poindester, A., Popovic, D., Primus, E., Pu, L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,

QY	1062	GAGAAATTACAATGAGCAGTGT	TTTGTGAATCTTGTATTCTTGC	ATTGTGCACTTGTGGAAGTGTG	1121
Db	131984	TAAAGATATCAATGACCAATGA	ATTAACTCCAGTTATTTATGT	TG-GTGTGGCTG	132042
QY	1122	TGTTGTCTGTGTGTCTGTGTG	TCTGTGTGTGTGTGTGTGTGTG	TGTGTGTGTGTGTGTGTGTGTG	1181
Db	132043	TGTGTGTGTGTGTGTGTGTGTG	TGTGTGTGTGTGTGTGTGTGTG	TGTGTGTGTGTGTGTGTGTGTG	132096
QY	1182	TGTCGTGTCTGTGTGTGTGTG	TCTCTCCCTCTTTTGATTTT	TGTGGCCTGGAATATT	1241
Db	132097	-----	TTTCCCTCTTTTCA	TTTGTGTGTGGAATATT	132132
QY	1242	TATTATTCATATTTCTTGAATG	GGGTAAACATCTTTAGATTGA	AGTTTCTCTCCTAG--	1299
Db	132133	TATATCTGTGTGTGTGTGTGTG	TCTTGTGGCAAGTTAACTCT	CTGGTGTAGAGTTTCTCTCTAGTA	132192
QY	1300	CTTTCTTTAGTCTGCATTTGA	AGATAGATATCTTTACATCTG	ATCTGATTTTATCTTAGAATG	1359
Db	132193	TCTTCTATAGGCTGGAATTTG	TGGATAAATATGTCTAAAT	TTGGTTTGTTCATGCAATA	132252
QY	1360	TCTTTCTTTCTCCAACTAAT	TGTACAGAAAGTTTCTTAAG	TGCAGTAGTCTGGCT--G	1417
Db	132253	TCTTATTTTCTCCACTATGA	TAAATGTAGTGTTTTGTCTG	GAATTAATAGTCTGGCTGAG	132312
QY	1418	ACATCTGTAGTCTCTTTGAG	TCTGTAGCACATCTGTGAGG	CCCTCTTACATTTTGAGT	1477
Db	132313	ACAATCTGTGGTCTCTCAG	ATCTGGAACATCTGCCAGG	CCCTCTTGACTTTTAGAGT	132372
QY	1478	TTCTATTGGRAAGTCAGGTG	TAAATCTTAAATACATCTG	CCCTTATATCTTAATTTGCTT	1537
Db	132373	CTCTGAT--GAGAAGTTGAG	TGTAATCTTAAAGGCTGCC	TTTATGTGTATTTCTGG--CAT	132430
QY	1538	TTTTCCCTTGGCATCTTTTAA	TATCTCTTTTGTCTATAC	TTTTPAGTGAATTTGATTA	1597
Db	132431	TTTTCCCTTTGCAGCTTTTAA	T--TCTTTCTTTGTCTG	TAAATTTAGTGTTTTGACTATT	132489
QY	1598	ATGCACCTGTGGGAGTTTCT	TTTTCGGTCCAACTATTTT	TGGTGTTTTGTATGCTTCTTGT	1657
Db	132490	AAGTGTGGAGGATTTCTTT	TCGTGATCCAACTCGTTG	TGGTGTTTTGTGAACTCTTGT	132549
QY	1658	ACCTTGATPAGCATCTCTTT	CTCAAGGTTAGAAATTTT	CTTTTGTG--GTTTCTTGAA	1716
Db	132550	ACAATTAAGCAGCTCCTT	CTTTTAGTTAGAGAAAT	TTCTCTATGACTTTTGTAAA	132609
QY	1717	AATATTTTCCCTGCTTTTGA	CGTG-----	CTTCTTCCCTTCTCTATTCCCTTT	1766
Db	132610	ACTATCTTCTGGCTTTTAA	GCTGGGAATCTTCTTCT	CCATTCACCTTACTCTTAGGT	132669
QY	1767	GGTTTTGCATAGTCTCTG	CTTCTGGATGTTTTATG	CGCTGGATTATTITGACATTA	1826
Db	132670	TGGTTTTTCTAGTGTCC	AGATTTCTCTGGATGTTTT	GTGTGACAGAACTTTCTAGATATA	132729
QY	1827	ACAATTTCTTTGACCAAG	TATCCATTTCTTCTATCT	TGTCTTCTACTGCTGAGATTTCTC	1886
Db	132730	ACAATTTCTTTGACTATG	TATCGAATTTCTCTATT	TGTACCTTCTACACCTGCGTTTCTC	132789
QY	1887	TCTTCTATCTCTTGTATT	CTGTGAGGCTGTCTCTG	AGGTTTCCGTGTGGGTTCTTA	1946
Db	132790	TCTCCCACTCTTGA-----	-----	TCTGTGTGTCTGTCTCTTCCCT	132829
QY	1947	ATTTTTTCATTTCCAGAT	TTTCTTCAGTTTGGGTTT	TGTTTATTAATTTCTATTCCACTT	2006
Db	132830	AGGTTTTCCATCTCCAG	GAATTTCTCAGTTTGT	TTTCTTTATTTCTATTTCCTATT	132889
QY	2007	TCAGTCTCTGAAATGTTTT	TACTCATTTTCCCTC--	--CCAGTATTTACATTTTTCATAGGTTT	2063
Db	132890	TGAAGTCTTGACGTGTTT	ATCTTTTCTTCTTCTT	CTTCACTCATTTGATGCAATTTCTCTATT	132949
QY	2064	CTTTAAATGGAATTAAT	CAATTTTCTCTTCAAGG	ACCTTTTATGAAATTCATATGATGT	2123
Db	132950	CTTTATGGGAATCAAT	CATGTCCTGTTTAAAG	ACCTCAATCGCTTCTCATGATTTGGAAT	133009

Qy	2124	TAAGTCTTGGCTTGCTGCTTCAGCTATGTTGCAATCTCAGGCGCT-----ATTG	2173
Db	133010	TAAGGTCAATTTTCTCGTGCTTGTGAAGACATCCAGGCTGCCATCATGAAGAG	133069
Qy	2174	TAATAGGGTTTTAGCAGGACATATTGCCCTGGTTGTTATGTCTGTGTTTTTGGCTTTGG	2233
Db	133070	AAGACTGGATTCTGTGTGCCATATTCCTGGCTCTGTACTGTGTCTTATGCTG	133129
Qy	2234	C 2234	
Db	133130	C 133130	
RESULT	28		
AL683804			
LOCUS	AL683804	198631 bp	DNA linear ROD 23-OCT-2002
DEFINITION	Mouse DNA sequence from clone RP23-17F2 on chromosome 1, complete sequence.		
ACCESSION	AL683804		
VERSION	AL683804.15	GI:24366485	
KEYWORDS	HTG.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Skuce, C.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Oct 24, 2002 this sequence version replaced gi:24210306.		
COMMENT			

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquerv@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi:, EMBL; Swi:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-17F2 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6

FEATURES	SOURCE
VECTID:10:Qualifiers	
location:1:15963	"Mus musculus"
osgbatism:"genomic DNA"	
db_xref:"Gene:10030"	
chromosome:"1"	
clone:"R223-1752"	
clone_id:"BC1-23"	

Query Match 30.9%; Score 704.103; DB 1; Length 198631;
Best Local Similarity 65.1%; Pred. No. 0.14;
Matches 1355; Conservative 0; Mismatches 574; Indels 153; Gaps 23

QY 270 TTCCCTCTCTCTTCCAAACACTCTATTTCTTGTGATTTCTATCTTGGCTCAATTTTAACTC 329
DB 186779 TTCACTGAATTTTGAAGAAGCTTAAATTTCTTTTATTTCTTCCCTAA-----TCATTG 186834
QY 330 AGTAGTGAGTTGTTGGTTTCCATAAGTTCTGAAGTTTCTGTTGTTTCTGTTGTTCTGTTG 389
DB 186835 AGTAGAGACCAATTCAGTTTCTCAATAGTAGGCTTTCTGTTGTTTCTGTTGTTGTTG 186894
QY 390 TTGTTATCTAGATTTAAGCTGTGTGTGTCAGATAGGA-CATPAGAGTATTTATTTCAATGT 448
DB 186895 AAG---TCCAGCTTTATTCATGGTGATCTGATAAGATCCATGGGTTTATTTCAATCTT 186951
QY 449 CTTTATCTGTGAGACTGCTTTGTTTGAATATGATATTCATTTTGGAGA--GTTTC 506
DB 186952 CTTGTATCTGTGGACTTGTCTTGTGACTGATTTGTTGGCAATTTTGGAGAAGGTTCC 187011
QY 507 ATAGGCTCTGACAGAAGGTACAGTCTTGTGTTTGGTGA---TAGTCTGTAATA 563
DB 187012 ATGAGGTCTGAGAAGAAGGTATGTTCTGTTTATTTGGTGAAATGTTCTATAGATGA 187071
QY 564 TCTCTAGTCTCACTTGTGTTTATGACATCAGTTAGTCTCAGCAATTTCTCTGTTCTGTTT 623
DB 187072 ATGTTAGGTTTCAATTTGAATATAACCTCTGTAGTTTCTGTTTATTTCTCTGTTAG-TTTC 187130
QY 624 TGTGAGATGACCTAACCTGTTGGAGAAATGGGTTATGAGTAGCCCACTATCTGTGTG 683
DB 187131 TGTCTTGAAGACATGTCATGTTGAAGTGGGTTGTGAGTCTCCCACTACTAATGTG 187190
QY 684 T-GAGGTCAATATGATTTAGCTGTAGCTGTGCTGTTTATGAACTTGGGTGACATT 742
DB 187191 TGGTGTCAATGTGTGATTTAAGCTTTAGTCAATGTTCTTTTACAAATGGGTGTCTT 187250
QY 743 GTGTTTGTGATGACATTAAGAAATGCAATGTCCTCTGTTGGTGA--TTTCTCTTTCATG 801
DB 187251 GCATTTGGGGATAGATGTTGAGAAATGGGATGACCTCTTAGTGGAATTTTCCCTTGA 187310
QY 802 CCTATGTAGTATTTCTCCCAATCTCATCTGCTTGTGTTTGGTTTAAAGTCTATTAG-TCA 860
DB 187311 AGTATGAAGTGTACTTCCCAATCTCTTTTGTGTTGTTGGTGAAGTCTATTTATTA 187370
QY 861 GATATTAATGACTGTATCGGCTGCTTCTTAGGGCAATTTGCTTAGAATATCTTTTCC 920
DB 187371 GTTATTGAATGGCTAATCCCAACTGTTTCTTGGGCTGTGTTGTTGAAACCTTTTCCC 187430
QY 921 ATCCCTTTTACT--CTAAGGTGATGTCATCCATGGTAGTGTCTTTTGGATGACGA 978
DB 187431 AGCCTTTATGTGACAGAGTACTATTTTGTGTTGATGATGTTCTTGTATAGCA 187490
QY 979 GTAGATGGATCTGTTTTCATATCCATCTGTATACCAAGTATCTTTTCTAGAGAAAT 1038
DB 187491 GAATGATGGATCTGTTTATGATCAATCCATCTGTATAGCCTTTGCTTTTATGTTGGAAAT 187550
QY 1039 AAGATCAATGATGATGTTGAGAAATTAATGAGCAGTGTGTTTGGATCTGTTT 1098
DB 187551 GAG-----TCTATTGAGTTTGAAGATGTTAATGACCAATGATTTGTTTCTCTGT- 187601
QY 1099 ATCTTGCACTGTGAAGT 1158
DB 187602 -----CTTTTGTGATATGTTGGTGTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 187655
QY 1159 GT 1218
DB 187656 GT 187707
QY 1219 TTGATTTTGGCCTGGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1278
DB 187708 -----GAATTTATTTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 187752
QY 1279 AGATTGAATTTTCTCTAGCCT--TCTTTAGTCTGATTTTGAAGATAGATTTCTTT 1336
DB 187753 GGGTTGGAGTTTCTCTTTTATTTATTTCTGTAGGCTGGGTAGTGGAAAGGTATTTT 187812
QY 1337 ACATCTGATTTTATCTTAGAATGCTTTCTTTCTCCAACTATTTGTGACAGAAAGTTTTC 1396

DB 187813 AAATTTGGTTTTGTTGGAAATCTTGATTTCTCCACTATGGTGACTGAAATTTGC 187872
QY 1397 TAAGTGCAGTAGTCTGGCTGACATCTCTAGTCTCTTGGAGTCTGTAGCACATCTGTGCA 1456
DB 187873 TGGGTGTAGTGTCTGTGACATCTGTGTTCTCTAGGTTTGCAGAACCTCTGCCA 187932
QY 1457 GGGCTTCTTACATTTGAGTTTCTATTGGAAGTCAAGTGTATTTCTATACATCTGC 1516
DB 187933 GGGCTTCTTGTGTTTGTAGAGTCTC-ATGTGATAAATGGGTATAAATTTGATAGTCTGC 187991
QY 1517 CTTTATATCTTAATTTGGTCTTTTCCCTTGGCATCTTTTAAATTTCTTTCTTTTGTCTAT 1576
DB 187992 CTTTATATGTTACTTGGAC-TTTTTCCTTGTCTGCTTTTAAAGTCTTTCTTTTCTGT 188050
QY 1577 ACTTTAGTATTTGATTTATTAATGACATGTGGGGAGTTTCTTTCCGTCCAACTATTT 1636
DB 188051 ATATTTAGTGTGTTGATTTATTTATGTTGGTGGAGGATTTTCTTTTGGTCCAACTGTT 188110
QY 1637 GGTGTTTGTATGCTCTTGTACCTTGTATAGCATCTCTTTCTCAAGTGTAGGAAATTT 1696
DB 188111 GGTGTTCTCTTACTTCTTGTACCTTTATAGCATCTCTTTCTTTAGTGTAGGTAATTT 188170
QY 1697 TCTTTTGTG-GTTTTCTTGAATAATTTTCCCTGCTTTTGACCTGC-CTTCTTCCCTTC 1754
DB 188171 TCTTCTATGATTTTGTGGAAGATGTTTCTCGGCTTTGAGTGCATAATTTCTCTTC 188230
QY 1755 CTCATTTC-----TTTGGTTTTGTAGTGTCTCTGGCTTCTCGGATGTT 1801
DB 188231 TTTCTATCTGTGTTATCTTAGTGTGTTTGTCTTTTCACTGTCCTCAATTTCTGATGTT 188290
QY 1802 TTATGCTCGGATTTATTTAGACTTAACATTTCTTTGACCAAGGTATCCATTTCTCTAT 1861
DB 188291 TTATGTCATGTAATTTTCAGATTTTGCATTTCTTTGACTGACATATTAATTTCTCTAT 188350
QY 1862 CTTGCTTCACTGCTGAGATCTCTCTCTTCTATCTCTTGTATCTCTGTTCTGTCAGTGAGCTGT 1921
DB 188351 TGTATCTCTATACTTGTAGGTC-----TCTATCGCTGTATCTGTTTGTGATCTTAT 188405
QY 1922 CTCGAGGTTCTGTTGGGTTCTTAATTTTTCAGATTTCTTCCATTTGAGTTGGGT 1981
DB 188406 GCTTTATCTCTGCTCTCTTCTTCCAGTTTATCTCCAGATTTGCTCAGTTGTGTT 188465
QY 1982 TTGTTTATTAATTTATTTCCACTTTCAGGCTCCTGAAATGTTTACTCATTTTCTCCCA 2041
DB 188466 TTTCTTATGCTTCTATTTTCCCTTTTTCAGGTTGTGACAGATTTTATCATATCTCTCACC 188525
QY 2042 GTATT----- 2046
DB 188526 TGTTTGATTTGATTTTCTGTATCTTATATCTTAGTGTCTTACTTTTAAATGCTCTAC 188585
QY 2047 -----TACATTTTTCATAGGTTTCTTTTAAATGATTTTATTCATTTCTCTTCAAGGACC 2098
DB 188586 TGTTTGAATGATTTTCTGTTATTTCTTAAAGATTTTACTCATGCTCTCTTAAAGGCT 188645
QY 2099 TTTTATGAATTCATAAATGATGTTAGTGTCTTCCCTTGTGCTTCACTGATGTTGAT 2158
DB 188646 TCTGTATCTTTTATAAATACATTAAGATCTCTCTCTGTTGTTTGTTCAT-TAGGTT 188704
QY 2159 TCTCAGGGCTATTTGTAATA-----GGGTTTTAGCAGGACATATTTGCTGTTGT 2210
DB 188705 ATCCAGGTTTGTGTAGGAGACAGCTGGACTCTGTTGGTATCATAGTCTCTCTGGGCT 188764
QY 2211 TATTTGCTGTGTTTTTGTCTTTGGCATATAGACGGCTGAGTTT 2252
DB 188765 TATTTGATTTGTCTGAAACTGTCTCTTTAGGAATCTGTTTT 188806

RESULT 29

AC101656/c

LOCUS

DEFINITION

267360 bp DNA linear HTG 23-APR-2003
Mus musculus clone RP23-306M24, WORKING DRAFT SEQUENCE, 47
unordered pieces.

ACCESSION

AC101656

VERSION

AC101656.2 GI:30018174

KEYWORDS

HTG; HTGS PHASE1; HTGS DRAFT.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

Biren, B., Nusbaum, C. and Lander, E.

JOURNAL

Mus musculus, clone RP23-306M24

REFERENCE

Unpublished

AUTHORS

2 (bases 1 to 267360)

AUTHORS

Biren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, K., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gerde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kellis, C., LaRocque, K., Lanazares, R., Landers, T., Lehotzky, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbak, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission

JOURNAL

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 267360)

AUTHORS

Biren, B., Nusbaum, C., Lander, E., Abouelkell, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbak, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

Direct Submission

JOURNAL

Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Apr 17, 2003 this sequence version replaced gi:17060431.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WtBR

Web site: <http://www-seq.wi.mit.edu>Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L17152

Center clone name: 306_M_24

----- NOTE: This is a 'working draft' sequence. It currently consists of 47 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 606: contig of 606 bp in length
 * 607 706: gap of 100 bp
 * 707 1376: contig of 670 bp in length
 * 1377 1476: gap of 100 bp
 * 1477 2112: contig of 636 bp in length
 * 2113 2212: gap of 100 bp
 * 2213 2934: contig of 722 bp in length
 * 2935 3034: gap of 100 bp
 * 3035 3825: contig of 791 bp in length
 * 3826 3925: gap of 100 bp
 * 3926 4857: contig of 932 bp in length
 * 4858 4957: gap of 100 bp
 * 4958 5869: contig of 912 bp in length
 * 5870 5969: gap of 100 bp
 * 5970 6546: contig of 577 bp in length
 * 6547 6646: gap of 100 bp
 * 6647 7092: contig of 446 bp in length
 * 7093 7192: gap of 100 bp
 * 7193 8443: contig of 951 bp in length
 * 8444 8243: gap of 100 bp
 * 8245 8856: contig of 613 bp in length
 * 8857 8956: gap of 100 bp
 * 8957 10016: contig of 1060 bp in length
 * 10017 10116: gap of 100 bp
 * 10117 10736: contig of 620 bp in length
 * 10737 10837: gap of 100 bp
 * 10838 11306: contig of 470 bp in length
 * 11307 11406: gap of 100 bp
 * 11407 12849: contig of 1443 bp in length
 * 12850 12949: gap of 100 bp
 * 12950 13716: contig of 767 bp in length
 * 13717 13816: gap of 100 bp
 * 13817 14302: contig of 486 bp in length
 * 14303 14402: gap of 100 bp
 * 14403 15778: contig of 1376 bp in length
 * 15779 15878: gap of 100 bp
 * 15879 16884: contig of 1006 bp in length
 * 16885 16984: gap of 100 bp
 * 16985 18208: contig of 1224 bp in length
 * 18209 18308: gap of 100 bp
 * 18309 19291: contig of 983 bp in length
 * 19292 19391: gap of 100 bp
 * 19392 20216: contig of 825 bp in length
 * 20217 20316: gap of 100 bp
 * 20317 21379: contig of 963 bp in length
 * 21379 21379: gap of 100 bp
 * 21380 22611: contig of 1232 bp in length
 * 22612 23836: contig of 1125 bp in length
 * 23837 23936: gap of 100 bp
 * 23937 24576: contig of 640 bp in length
 * 24577 24676: gap of 100 bp
 * 24677 25451: contig of 775 bp in length
 * 25452 25551: gap of 100 bp
 * 25552 26907: contig of 1356 bp in length
 * 26908 27008: contig of 100 bp
 * 27008 28014: contig of 1007 bp in length
 * 28015 28114: gap of 100 bp
 * 28115 29141: contig of 1027 bp in length
 * 29142 29241: gap of 100 bp
 * 29242 30593: contig of 1352 bp in length
 * 30594 30693: gap of 100 bp
 * 30694 31784: contig of 1091 bp in length
 * 31785 31884: gap of 100 bp

* 31885	33263: contig of 1379 bp in length	Db	125647	ATGAGTGTCTGAGAAGAGGTATGTTGGTGTATTTGGGTGAAATGTTCTATAGATGTA	125588
* 33264	33363: gap of 100 bp				
* 33364	34467: contig of 1104 bp in length	QY	564	TCTCTAGGTCACCTTGGTTTATGACATCAGTTAGTCCAGCATTTCTCTGTTGGTTTTT	623
* 34468	34567: gap of 100 bp				
* 34568	35942: contig of 1375 bp in length	Db	125587	ATGTTAGGTTCAATTTGAATCAATCACTCTGTTAGTTCTGTTATTTCTGTTAG-TTTC	125529
* 35943	36042: gap of 100 bp				
* 36043	37569: contig of 1527 bp in length	QY	624	TGTTGAGATGACCTAACTGTTGGAGAGATGGGGTATTGAAGTAGCCACATCTGTGTG	683
* 37570	37669: gap of 100 bp				
* 37670	39143: contig of 1474 bp in length	Db	125528	TGTTCTTGATGACATGCCAATGTTAAGAGTGGGGTGTGAGTCTCCACACATAAATGTG	125469
* 39144	39243: gap of 100 bp				
* 39244	41309: contig of 2066 bp in length	QY	684	T-CAGGTCAATATGTGATTTAGCTGTAGCTGTCTGTTTATGAACTCTGGGTGACAT	742
* 41310	41409: gap of 100 bp				
* 41410	42879: contig of 1470 bp in length	Db	125468	TGGTGTTCATGCTGTAATTAAGCTTTAGTCATGTTCTTTTACAAATGGGGTCTCT	125409
* 42880	42979: gap of 100 bp				
* 42980	44743: contig of 1764 bp in length	QY	743	GTGTTGGTGCATAGACATTAAGAAATTCGAATGCTCTCTGGTGA-TTTTCCCTTTGATG	801
* 44744	44843: gap of 100 bp				
* 44844	46754: contig of 1911 bp in length	Db	125408	GCATTTGGGSCATAGATGTTTCAGAAATGGGATGACCTCTTAGTGGATTTTTCCTTTGATA	125349
* 46755	46854: gap of 100 bp				
* 46855	48072: contig of 1218 bp in length	QY	802	CCTATGTAGTATCTTCCCAATCTCATCTGCTTGTAGTTTGGGTTTAACTCTATTAG-TCA	860
* 48073	48172: gap of 100 bp				
* 48173	50344: contig of 2172 bp in length	Db	125348	AGTATGAAGTGTACTTCCCAATCTCTTTGATGCTTTTGGTGAAGAGCTCTATTATTATA	125289
* 50345	50444: gap of 100 bp				
* 50445	53613: contig of 3169 bp in length	QY	861	GATATTAAATGACTGTATCGGCTTGTCTTTPAGGSCCATTTGCTTGAATAATCTTTTCC	920
* 53614	53713: gap of 100 bp				
* 53714	94770: contig of 41057 bp in length	Db	125288	GTATTAGAATGGCTAAATCCAACTTGTCTCTGGGCTGTGTTGTTGTAAGAACCTTTTCCC	125229
* 94771	94870: gap of 100 bp				
* 94871	178917: contig of 82047 bp in length	QY	921	ATCCCTTTTACT--CTAAGGTGATGCTATCCATGCTAGGTGCTCTTTTGGATCAGCA	978
* 178918	179017: gap of 100 bp				
* 179018	267360: contig of 90343 bp in length.	Db	125228	AGCCCTTTATGTTGAGACAGTGAATTTTGTGTTGAGATGTGTTTCTGTATAGAGCA	125169
FEATURES					
Location/Qualifiers					
1..267360	/organism="Mus musculus"	QY	979	GTAGATGGAATCTGTTTTCATATCCATTCGTATACCCAGTATCTTTTCTAGAGAAAT	1038
1..606	/no_type="genomic DNA"				
707..1376	/db_xref="taxon:10090"	Db	125168	GAATGATGGAATCTGTTTATGCAATCCATTCGTAGCCCTTGTCTTTTATGTTGGAAAT	125109
1..606	/clone="RP23-306M24"				
1..606	/clone_lib="RPC1-23 Female Mouse BAC"	QY	1039	AGATCAATGAGTCATGATGTTGAGAAATATCAATGACGAGTGTGTTGGAATCTTTGT	1098
707..1376	/note="assembly_fragment"	Db	125108	GAG-----TCTATTGATGTTAAGAGATGTTAATGACCAATGATGTTACTTCTCTGT-	125058
1477..2112	/note="assembly_fragment"	QY	1099	ATCTTGCACCTGTGAAGT	1158
2213..2934	/note="assembly_fragment"	Db	125057	-----CTTTTGTGATATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	125004
3035..3825	/note="assembly_fragment"	QY	1159	GT	1218
3926..4857	/note="assembly_fragment"	Db	125003	GT	124952
4958..5869	/note="assembly_fragment"	QY	1219	TTGATTTTGGCCCTGGAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	1278
5970..6546	/note="assembly_fragment"	Db	124951	-----GAATTTATTTCTTGT	124907
30.9%; Score 704.103; DB 1; Length 267360;	Query Match	QY	1279	AGATTGAAGTTTTTCTCTAGCCT--TCTTTAGGCTGCAATTTGAAGATAGATATCTTT	1336
65.1%; Pred. No. 0.18; Indels 153; Gaps 23;	Best Local Similarity	Db	124906	GGGTTGGAGTTTTTCTCTTTTATGATTTCTCTGTAGGCTGGGTAGTGGGAAGTATTTT	124847
0; Mismatches 574;	Matches 1355; Conservative	QY	1337	ACATCTGATTTTATCTAGAAATCTTTCTCTCAACTATGTCAGAGAAAGTTTTTC	1396
		Db	124846	AAATTTGGTTTTGTGTGGAATATCTTGATTTCTCCACTATGGTGACTGAAATTTGC	124787
270 TTCCCTCTCTCTCCAAACACTCTATTCTTTGATTTCTATCTATCTTGGCTCAATTTAACTC	329	QY	1397	TAAGTCAGTAGTCTGGCCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCA	1456
TTCACTGAATTTTGAAGAGTCTTAAATTTCTTTCTTTATTTCTTCCCTAA-----TCATTG	125825	Db	124786	TGGGTGTAGTAGTCTGTGCTGACATCTGTGTTCTCTTAGGGTTTCAAGACCTCTGCCA	124727
330 AGTAGTAGTGTGTTGGTTTCCATAGTTTCTGATTTCTGTTGTTCTGTTGTTGTTGTTG	389	QY	1457	GGGCCCTCTTACATTTGAGTTTCTATTGGAAAGTCAAGGTAGTCTTAATACATCTGC	1516
AGTAGAGGCCATTCAGTTTCTCAATATATGAGCTTTCTGTTGTTCTGTTGTTGTTGTTG	125765	Db	124726	GGCCCTCTTCTGTTTTTAGAGTCTC-ATGTGATAAATGGGTATAAATTTGATAGTCTGC	124668
390 TGTATCTAGATTTAAGCTGTGGTGCAGATAGGA-CATAGAGATTTATTTCAATGT	448	QY	1517	CTTTATGCTTAATGTTGTTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTAT	1576
AAG---TCCAGCTTTATTTCCATGTTGATCTGATAAGATCCATGGGTTTATTTCAATCTT	125708	Db	124667	CTTTATGCTTAATGTTGTTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTAT	124609
449 CTTTATCTGCGAGACTGTTGTTTGGAAATGATGATTTCAATTTGGAGA--GTTTC	506	QY	1577	ACTTTTAGTCAATTTGATTTATTAATGACCTGTGGGAGTTTCTTTTCCGTCCTAATCTATT	1636
CTTTATCTGTTGGACTTGCTTTGTGACTGATTTGGGCAATTTGGAGAGGTTCC	125648	Db			
507 ATAGGCTGCTGACAGAGGTTACAGTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	563	QY			

124608	ATATTTAGTGTGTTTCAATTAATTAATGGTGGGAGATTTTCTTTTGGTCCAAATCTGTTT	124549
QY	GGTGTTTGTATGCTCTCTGTACCTTGATAGGCAATCTCTTCTCAAGGTTAGGAAATTTT	1696
124548	GGTGTTGCTTACTCTTGTACCTTTATAGGCAATCTCTTCTTAGGTTAGGGAATTT	124489
QY	TCCTTTTGG-GTTTTCTTGAAAATATTTTCCCTGCTTTGACCTGC-CTTCTTCCCTTC	1754
124488	TCCTTAAGATTTTGTGAAAGATTTTCTCGGCCCTTTGAGCTGCAAAATATTTCTCCTTC	124429
QY	CTCTATTCC-----TTTGGTTTTGATAGTGTCTCTGGCTTCCCTGGATGTT	1801
124428	TTCTATTCTGTTATTCTTAGGTTTGGTCTTTTCCACTGTGCCAAATTTTCTGGATGT	124369
QY	TTATGCTCGGATTATTTTAGACTTAAATTTTCTTTGACCAGAGGTATCCATTTCTTCTAT	1861
124368	TTATGTCATGTAATTTTCAGATTTTTCAGATTTTCTTGACTGACATATAATTTCTTCTAT	124309
QY	CTTGCTTTCAGTCCGTGAGATCTCTCTTCTATCTCTTGTATTCTGTCACTGAGGCTTCT	1921
124308	TGTATCTTCTATACTTGAGATC-----TCTATCGCTGTATCTGTTGTGATACTTAT	124254
QY	CTTTCAGGTTCTCTGTGGTTCCTAAATTTTCTTATTTCCAGATTTCTCTCAGTTTGGGTT	1981
124253	GGCTTTATCT	124194
QY	TTGCTTATTAATCTTATTTTCCACTTTCAGGTCCTGAAATGTTTTACTCATTTTCTCCCA	2041
124193	TTCTTTATTGCTTCTATTTTCCCTTTTCAGGTGTGGACAGTTTTATTCATATCCTTACC	124134
QY	GTATT-----GTATT-----GTATT-----GTATT-----GTATT-----	2046
124133	TGTTTGATGTATTTTCTGTATCTTATACTTAGTTGCTTACTTTTTTAATGTCCTACC	124074
QY	-----TACATTTTCATAGTTTCTTTAATGGAATTAATTCATTTCTCTTCAAGGACC	2098
124073	TGTTTGAATGATTTTCTGTATTTCTCTTAAGAGATTTACTCATGTCCTCTTAAAGGCT	124014
QY	TTTTTANGAATTCATAAAATGATGTTAAGTTCCTTGCTTGTCTTCAAGCTATGTTGCAAT	2158
124013	TCTGTCTATCTTTATAAAATTAACATTTAAGATTCCTCTTCTTGTACTTGTTTTCAT-TAGGGT	123955
QY	TCTCAGGCGCTATTCTGAATA-----GGGTTTTAGCAGGACATATGCTCTCTGTTCT	2210
123954	ATCCAGGTTTGTGTAGAGGACAGCTGGACTCTGTGTATCATAGTCTCTCTGGTCT	123895
QY	TATGTCTGTGTTTTTGTGTTTGGCATATAGACGGCTGAGTTT	2252
123894	TATTCATGTTGTGTTCTGAAACGTCTTTTAGGAATCTGGTTTT	123853

Db	117069	TCCTATTTATTTTTCGACCCAGTGGTCA*TGATGAAGGGGTATTACAGTTTCAATGAGT	117128
Qy	358	TTGTAAAGTTTTCTGTGTGTTCTGTGTGTGTGTG*TATCTAGATATTAAGCTGTGCTGGT	417
Db	117129	TTGTAGGAATTTCTGT*GTTTCTGATATAAAGGT-----CAGCTTTAATCCATGCTGGT	117182
Qy	418	CAGATAGGACATAGAGATTAATTTCAATTCCTCTTTTATCTGTGAGAGACTTTCCTTTCGTTTT	477
Db	117183	CAGATAAAGATACATGGGGGTATTATTCATTTTTATATGTGTGTGAGTTTGTCTTCAATGAC	117242
Qy	478	GAATATATCTAATTCATTTTGGAGA--GTTTTCATAGGGTCTGCACAAGAAGGTACACATC-	533
Db	117243	TAAGTATATGATCAATGTTGGAGAAGGTTTTTATGAGGTGATGAGAAGAAGGTATATTCT	117302
Qy	534	TTTGTGTTTTGGTGAATAGTCGTGAATATCTCTAGGTCCACTTGGTTTATGACATCATG	593
Db	117303	TTTGTGTTTTGGTGAATGTTCTGTACATGCTGTTAGATCCATTGTTTCATACATCTTA	117362
Qy	594	TATAGCTCCAGCATTTCTCTGTGTTGGTTTTTTTGTGATGACCTAACTGTGTGAGAGAAAT	653
Db	117363	TTAGTTTCATTATTTCTCTGTTTACTTT-----CTGGCTGGATGAACCTGAG	117408

Db	118430	ATGGACCATGAGTTGGGAATCTTCTCTTCTATGCCATTAATCTTAGGTAGGTTTT	118489
Qy	1769	TTTTTGCAATGTCCTGGCTTCTCGATGTTTTATGCTCGGATTAATTTAGACTTAAC	1828
Db	118490	TTCTGGCACATATCCACAGATTCCTGGATGTTTGTATCAGGAAGTTTATGATTAAC	118549
Qy	1829	ATTTCTTTTGACCAAGGTATCAATTTCTCTATCTTCTCTCTCTGCTGAGATTCCTTC	1888
Db	118550	ATTTCTTTTGACAGATGATCAATTTCTTCTCTGATATCTTACACATGAGGGTCTCTC	118609
Qy	1889	TTCTATCTCTGATTTCTGTCAGTCAGCTTGCTCTGAGGTTCTGTGGGTT-CTTAA	1947
Db	118610	TTTTATCTCTGATTCCTGTTGGTGATCTCTATCTGTGTTCTGTCTCTTACCAAG	118669
Qy	1948	TTTTTTCATTTCCAGATTTCTCTTCTGATTTGGGTTTTCTTATTAATCTATTTCCACTTT	2007
Db	118670	GTTTTCATCTCCAGATTCCTCAATTTGTTGTTTCAATTTACTGCTCTTATTTCCATTTT	118729
Qy	2008	CAGGTCCTGAAATGTTTACTCAT- - -TTTCTCTCCAGTATTTACATTTTCATAGTTTC	2064
Db	118730	TATGCTTTGACAGTTTTACTTCATTTCTTTCACCTGTTGATGTTATGTCCTGTATTC	118789
Qy	2065	TTTTATGGAATTAATTCATTTCTCTTCAAGACCTTTATGAATTCATAAATGTAATGTT	2124
Db	118790	TTTAAACGGAATTAATCAATTTCTCTTTAAAGGCTCTATCATTTTCATAGTTTTCATAT	118849
Qy	2125	AAGGTCCTGCTGCTGCTTCAGCTATGTT- -GCATTTCTCAGGCTCATTTGTAATAGGTT	2183
Db	118850	AAGTCAATTTCTGCTGCTTCAGCTGTGTTAGGTACTCAGGGCTGCTGACCAAAGA	118909
Qy	2184	TTAGCAGGACATATTTGTCCTGGTTGTTATGTC- -TGTGTTTGTCTTGGCATATAGAC	2242
Db	118910	GCTAGGCTCGTGATAACGTGGCTGATGTTGACTTATGTTCTTACTCTGGCTTTAGCC	118969
Qy	2243	GGCTGAG- -TTTGGGATGATTTGAATCTCAGGTCGTGAT	2279
Db	118970	TTCTGGTTTTCATGCTTATTTAGGTTTAGATGCTGTT	119007
RESULT 31			
AC126499/c			
LOCUS	AC126499	211245 bp	DNA linear HTG 20-NOV-2002
DEFINITION	Rattus norvegicus clone CH230-202L20, WORKING DRAFT SEQUENCE, 8 unordered pieces.		
ACCESSION	AC126499		
VERSION	AC126499.4	GI:25138312	
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE 1	(bases 1 to 211245)		
AUTHORS	Muzny, D., Marie, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Auguano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenship, K., Blyth, P., Brown, M., Bryant, N., Buahy, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dirh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorjse, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, K., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hughes, M.,		

RESULT 31
AC126449/c
LOCUS
DEFINITION
AC126499 211245 bp DNA linear HTG 20-NOV-2002
Rattus norvegicus clone CH230-202L20, WORKING DRAFT SEQUENCE, 8
unordered pieces.
AC126499 4 GI-25138312
AC126499 HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLLTOP.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
References 1 to 211245)
Muzny, D., Marjé, M., Merzker, M., Lee, A., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhaq, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
David, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C. M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Georgegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, N., Hughes, M.,

QY	648	GAGAAATGGGGTATTGAAGTAGCCCACTATCTGTGTGA--GGTCAATATGTGATTTTACG	706
Db	122509	GAGATCGGATATCAAGTCTCACACTATTACTGTATGAGGTCGAGATGTGATTTAAGC	122568
QY	707	TGTAGCTGTGCTGTTTATGAACCTTGGTGACATTTGTGTTGGTGATAGACATTAAAG	766
Db	122569	TGTAGTAGCATTTGTTTACAACCTTTGGTGCTTTTGTGTTTGGGACATAGATATCAATG	122628
QY	767	ATTGCAATCTCCTCTTGGTGA--TTTTCTTTGATGCCTATGTAGTATCTTCCCAATCT	825
Db	122629	ATTGACATGGCATCTTGGTGGATTTTACTATGATGAGATATAGTGTCTCTCCCAATCT	122688
QY	826	CATCTCTTAGTTTTGGGTTTAAGTCTAATTA--GTACAGATATTAATAAGACTGTATCGGT	884
Db	122689	CCTTTGATTAGTTTTAGTTTGAAGTCTATTTTGGCAGATATTTAAAGGCTATAGACCT	122748
QY	885	TGCTTCTTAAAGGCCATTTGCTTAGAATATCTTTTCCATCTTTTACTCTTAAGGTGATGTC	944
Db	122749	TGTTACCAGCTTGCTTCTCATGTCCAATTTGCTCATATACCTTTGCCAACCCAAATTT	122808
QY	945	TATCCATGTAGGTTGTCTTTTGGATGACAGAGTAGGATAGATCTTGTTTCAATACC	1004
Db	122809	TCTTGATGTGAACTGTGTCTTCTGGTTGCAGGAGAAGAT--GATCTGGTTTTCTCATCC	122867
QY	1005	ATTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATTGAGTC--ATTGATGTGA	1063
Db	122868	ATTCTGTACTTGTTACTTTTGTGTGCTTTTGTCTGGGGAACTGGGACTATTAATATGA	122927
QY	1064	GAATTATCAATGAGCAGTGTTTTGGATCTTGTTATCTTTGCACCTTGTGAAGTGTGTGTG	1123
Db	122928	GATATATCAATCACTCATGTTTGATTATGATTTCTGTATGTTTGTGTTGTGTTG	122987
QY	1124	TG	1183
Db	122988	TG	123047
QY	1184	CTGTGTCTTGTGTGTGTGTCTCCCTCTTTTGTGATTTTGTGGCTGGAAATATTTA	1243
Db	123048	TGTGTG--TGTGTGTGTGTGTCTTACTTTCTTTTCGGTGTGTGTGAGATATTAAT	123104
QY	1244	TTATTCATATTTCTTGAATGTGGGTACACTTTTAGATTGAAGTTTTTCTCCTA--GGC	1301
Db	123105	TTCCCTGTGATTTCAAGTTTGGAAATTAACCTCTTAAGTTGGAGATTTCTTCTATCAC	123164
QY	1302	TTCTTTAGTCTGCATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAAGATGTC	1361
Db	123165	TTCTGCAGAGCTGGATTTGTAGATAACTATTTGTCTAAATTTGACTTTTATCATGGAATATC	123224
QY	1362	TTTCTTTCTCCAACTATTGTGACAGAAAGTTTTCTTAAGTGCAGTACTCTGGCTCAGCAT	1421
Db	123225	-----TTCTCTATAATGGTGTAGGAAGTTTTGCTGGAATAGTAGGTTGAGCTGGCAT	123279
QY	1422	CTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCCCTTTACATTTGAGTTTCT	1481
Db	123280	CTGTGGTCTCTTAGAATCTGCAGCAAACTTTTCAGGCCCTTTCTGGCTTTTAAAGTCT-T	123338
QY	1482	ATTGGAAGAGTCAGGTGAATTTCTAATACATCTGCCTTTATATGTTAATTTGGTCTTTTTT	1541
Db	123339	ATTTGAAGAGTCAGTTATTAATTTCTATATGTCTGCATTTATATGTTACTTCTCTTCT	123398
QY	1542	CCCTTGCACTTTTAAATATCTTCTTGTCTATACCTTTTAGTGATTTGATTTATATGC	1601
Db	123399	-CCTTCAGCTTTTAAATA-----TTTTTTTATTCTGTGTTTGTAGTATTTTGACCAATATG	123453
QY	1602	ACGTGGGAGTTCTCTTTCGGTCCCAATCTATTTGGTGTGTTGTGTATGCTCTTTGTACC-	1660
Db	123454	TCAAT--AGGACTTTCTTCTCGGTCCAATCTATTGGCAATCTGTATATCTCTGTACCT	123512
QY	1661	TTGATAGGCATCTCTTTCTCAAGGTTAGGAAATTTTCTTTTTTGGTTTTCTTGAAATA	1720
Db	123513	TTAATAGGCATCTCTTCTAATAGGTTAAGAAA-----ATCTATGATTTTGTGTAATAAT	123566

Qy	1721	TTTTCCTCCGCTTTTGACCT - GCCTTCTTCCCTCCCTCTCTATTCCTCTTGGTTTTGCAATAG	1779
Db	123567	ATTTCCTGGCGTTTGGACCTGGGGTTTCTTCTCTCTTGAATTCCTATTAATCTTTGGTTT	123626
Qy	1780	TGCTCTGGCT - TCCTGGATGTTTATAGCTGATTAATTTAGACTTAACATTTTCCTTT	1837
Db	123627	GGTCCTTTCTTAGTCCCGAGATTTCGCAATGACATTAATTTTA - ATTACATTTTCCTTT	123685
Qy	1838	GACCAAGGTATCCATTTCTTCT - ATCTTGCTTCACTGCTCGAGATTCCTCTCTATC	1895
Db	123686	TACCCATGAATCTTTTCTTCTCTCTGTTGATCTTCAACTCTTGGTCTCTTCTCTCTGTC	123745
Qy	1896	TCTTGTATTCCTGTCAGTCAGGCTTGTCTGAGGTTCCCTG - TTGGGTTCTTAATTTTTC	1954
Db	123746	AATTGTATTCCTGTTGGTGAAGCTTGCCTCTGTAGTACTGTTTGAATTCCTAAATTTTC	123805
Qy	1955	ATTTCAGATTTCCCTTCAGTTTGGGTTTGTATTATTAATTTCTATTTCCTCTTCAGGCTC	2014
Db	123806	ACTTCTATAATCCCTCGTTTGGATCTCCTTTATTGATCTATTTCCTATTTTAAGTCT	123865
Qy	2015	TGAATGTTTACTCATTTTCTCCAGATTAATACATTTTTCATAGGTTTCTTTAAATGGAT	2074
Db	123866	TAAACAGTTTATTCA - TTTGCTTACACTGTTGTGTTTTTCCCTGTTTCTTTAAAGGAT	123924
Qy	2075	TTATTCAATTCCTCTTCAAGGACCTTTTATGAATTCATAAAATGTTATTTAAGGCTCCTTG	2134
Db	123925	TTATTCAATTCCTCTTTAAGGA - ATTGCTATCTTCATAAAGGCTCTTTTAAGGCTCTTT	123983
Qy	2135	CTTGTGCTCAGCTAATGTCATCTCAGGGCTATTGTAATAGGGTTTTCAGGAGAC	2194
Db	123984	ACTTGGGCTTCTTAT - -----GGTAGGGTTGCTGAGCTTTTAGTGAGAC	124026
Qy	2195	ATATTGTCCTGTTGTTATTCCTGCTGTTTTGCT - TTGCAATATAGCGCTGAGTTTG	2253
Db	124027	ATAGTACCTGCGCTTTATGATTATCTTTTTTATGCTAGTCTTAGGCATGTGGATAG	124086
Qy	2254	GGATGATTGTAATTTAGTGCTGA	2278
Db	124087	AGATGATTATACATCTAGTGCTGA	124111
RESULT 33			
AL669976			
LOCUS	Mouse DNA sequence from clone RP23-15J3 on chromosome X, complete sequence.	227762 bp	linear
DEFINITION			ROD 24-MAY-2002
ACCESSION	AL669976		
VERSION	AL669976.11	GI:21213362	
KEYWORDS	HTG.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 227762)		
AUTHORS	Brown, J.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: hamquery@sanger.ac.uk		
COMMENT	On May 25, 2002 this sequence version replaced gi:21068533. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e. phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following		

abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Rp23-1503 is from the RpCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES		Location/Qualifiers	
source	1..227762		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10090"		
	/chromosome="X"		
VECTOR: pBACe3.6.		/clone_lib="RPECI-23"	
Query Match 30.7%; Score 699.902; DB 1; Length 227762;			
Best Local Similarity 69.1%; Pred. No. 0.16; 511; Indels 59; Gaps 24;			
Matches 1276; Conservative 0; Mismatches 511; Indels 59; Gaps 24;			
QY	294	TATTTCTTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAGTTTGGTTTCCAT	353
Db	168005	TATTTTCAATTTTCCCTGACTGAGCAATCATTTGAGTAGAGAGTTGTTCAAGTTCCAA	168064
QY	354	AAGTTTCTAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	413
Db	168065	GAGTTTGAAGCTTCTGTTGTTTC---TGTGCTGTGTAATCCAGCTTTAATCTACAG	168121
QY	414	TGGTCAGATAGACATAGAGTATTTTCAATTTCTTTTATCTGTCGAGACTTGTCTTG	473
Db	168122	TAGTCTGATGAGTACAGGGGTTGTTTCAATTTTCTTGTCTGTTTGAGACTTGCTCT	168181
QY	474	TTTTGAAATATGATTCATTTTGGAG--AGTTTCATAGGGTGTCTGCAAGAAGGTACAG	531
Db	168182	TATTTGAGTATGTGTCACACTTTGAGAAAGTGTCATGAGTTGCTGAGAAGAAGGTCCAT	168241
QY	532	TCT--TTGTTGTTTGGTGAAATAGTCTGTAATATCT--CTAGGTCCACTTGGTTTATGACA	589
Db	168242	TATATTGCAATGGGTGAATGTTTGGTAATATCTGTTATGTTTATTAGTTTATTACT	168301
QY	590	TCAGTTAGCTCCAGCATTTCTCTGTTTGGTTTGTGATGATGACCTAACTGTTGGAGA	649
Db	168302	CCAATTAGCTCCAGCATTTCTCTGTTTGTCT---AGACGATTTGCCATTGGTGA	168356
QY	650	GAATGGGTATGAGTAGCCCACTATCTGTGTGTA--GGTCAATATGATGATTTAGCTG	708
Db	168357	GAATGGGGTATTAACCTCCCTACTATCAATGTGTTACGATCAATGTGATTTAACTG	168416
QY	709	TAGCTGTGCTTGTTTTATGAACCTGGGTGACATTTGTTGGTGCATAGACATTAAGAAT	768
Db	168417	TGGTAGTTTTTGTTTTACAGTGTGGGTGCCCTTGTGTTGGAGCATAGATTTAGAAT	168476
QY	769	TGCATGTCCTCTGGTGGGA--TTTTCTTTGATGCGCTATGATGATTTCTTCCCAATCTCA	827
Db	168477	TAAATATCATCTTGATGGATTTTCTCTCATGAATATGATGATGATGCTATGCTCCCATCTC-	168535
QY	828	TCTCCTTAGTTTTCGGTTTAACTATTAGTTCAGATTAATAAATGACTGTATCGGCTTGC	887
Db	168536	TTTCTTAATTTGCGTTTGA--TATTTGTTAGATTAATAAATGGGTACACCGTCTTGC	168593
QY	888	TTCTTAGGGCCATTTGTTAGAAAT--CTTTTCCATCTTTTACTCTAAGGTGATGCTA	946
Db	168594	TTCTAAGGTCCATTTGGCATGAAATATCTTTTCCCAACCAATTTACCTTTGAGTAAATGTTTA	168653
QY	947	TCC---ATGGTAGGTTGTCTTTTGGATGCGAGTAGGATGGATCTGTTTTCATATC	1003
Db	168654	TACTTGATGTTGAGGTATGTTTCTCATATCCAGTAAAGGATGGGTTCTGTTTCTGTACC	168713
QY	1004	CATTTCTGTTACCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGTTGA	1063
Db	168714	CATTTCTGTTAATCTGTGCTTTTATTTGGGAATGA-----GGCCATTAGTAATGA	168764

Qy	1064	GAATTATCAATGAGCAGTGTGTTGTGATTTCTTTTATCTTGCACTTGAAAGTGTGTG	1123
Db	168765	AAAGTATCAATGACAATGATTTAAATTTCTGCTATTCTTACGATGTTGTGGCAATG	168824
Qy	1124	TGTTG	1183
Db	168825	TG	168884
Qy	1184	TCTG	1243
Db	168885	AGAGAGAGAGG-----GAGAGAGTCTCTTACTTTGGTTTGTCTGATGTGTGTTATTTA	168938
Qy	1244	TTATTCAATATTTCTTGAATGTGGGTACATCTTTAGATTGAAGTTTCTCTCTA--GCC	1301
Db	168939	TTCTCTGTGTTTTCATGGAATGTAGTAACCTCTTGGGTAGAAATTTTCTTCTTACACT	168998
Qy	1302	TTCTTTAGTCTGCAATTTGAAGATAGATATCTTTTACATCTCATTTTATCTTTAGAAATGTC	1361
Db	168999	TTCTGTAGACTGGAATTTGTAGATGATATTCCTTAATTTGACTTTTTCGTGGAATATA	169058
Qy	1362	TTCTTTCTCCAACTATTGTGTGACAGAAAGTTTCTTAAAGTGCAGTAGTGTGCGCTGACAT	1421
Db	169059	T---TTTCTCTATTGTGTGTGA-----TTTGTCTGAGTATTGTAGTCTGACTGGCAT	169108
Qy	1422	CTGTAGTCTCTTGGAGTGTGTAGCACATCTGTGCGAGGCTCTTACATTTTGTAGTTTCT	1481
Db	169109	CTGTAGTCTCTTAGAATCTGTAGAAATATCTGTCCATGCCCTTCTAGCTTTTAGAGTCTCT	169168
Qy	1482	ATTGGAAGTCCAGGTGTAATTTCTAATACATCTGCTCTTTATATGTTAAATGTGCTTTTT	1541
Db	169169	GTT--GAGAAGTCCAGGCATAATTTCTAATAGTCTGTGTTTATATGTTACTTGTCTCT--TT	169226
Qy	1542	CCCTTGCATCTTTTAAATATCTTCTTCTTCTTCTATCTATCTTTTGTAGTGTGATTAATGCG	1601
Db	169227	CCCTTGCAGGTTTAAACATCTTCTTCTTCTGCTGACATTTAGTGTGTTTGGTTATTTT	169286
Qy	1602	ACTGTGGGAGTTTCTTTCCGGTCCAACTATTTTGGTGTGTTTGTATGCTTCTTGTACCT	1661
Db	169287	GATGGAGGA--TTTTTTCTGCTCAATCTATTTGGTCTCTGTAAAGCTTCTTCTACT	169345
Qy	1662	TGATAGCATCTCTTTCTCAAGGTGAGAAATTTTCTTTTGGTTTGGTTTCTTTG--AAAATA	1720
Db	169346	TCATAGGCATC---TTCCTTAGGTTAGGAAATTTTCTTGTAGGATTTTGTGAAAAATA	169402
Qy	1721	TTTTCCCTGCTTTTGACCTGCC--TTCTTCCCTCTCTCTATCTTCTTGGTT--TTTGCAT	1777
Db	169403	TTTTGAGACTTTTGAGCTGGGATTTCTTCTTCTTCTCAATTTCCCTAGTAACTTGGTTT	169462
Qy	1778	AGTGTCTCTGGCTTCTGTGATGTTTATGCTGGAATTTATTTAGACTTAAACATTTTCTTT	1837
Db	169463	GGCTTCCCATATTTCTCTGGATACTTTTGTCTCAGGAAATTTTTCACAAATAACA--TTTCTTT	169521
Qy	1838	GACCAAGGTATCCATTTCTTCTTCTTCTGCTTCACTGCTGAGATTTCTTCTTCTATCTC	1897
Db	169522	GACCAAGGTATCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	169581
Qy	1898	TTGATTTCTGTCAGTGGCTGTCTGAGGTTCTCTGTTGGTGTCTTAAATTTTCTTCAAT	1957
Db	169582	TTGATTTCTGTTGGTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	169641
Qy	1958	TCAGATTTCTTCTGATTTGGTTTGTGTTTATTAATTTCTTATTTTCACTTTTCAAGTCTTGA	2017
Db	169642	TCAGAAATTAACCTGCTGTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	169701
Qy	2018	AATGTTTCTACTCATTTTCTCTCCAGTATTTTCAATTTTCTATAGTGTCTTCTTAAATGGAATTA	2077
Db	169702	ACAGTTTATTTGTATGGGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	169761
Qy	2078	TTCTTTCTTCTTCAAGGACCTTTTATGAAATTCATAAAATGATGT	2123
Db	169762	TTGTTTTTCTTAGATTAATAAAATTTATTCATCAATTTAAATTTTATTTT	169807

QY 450 TTTTATCTGTGGAGACTTGTCTTTGTTTGAATAATATATCAATTTTGGAGAGTTTCATA 509
Db 206083 TTGTGTCTGTGGAGCTTGATCT-----TTGGAAAAGTTTCAIG 206044
QY 510 GGGTCTGACAGAAGAGTACAGTCTTTGTGTGTTTGGTGAATAGTCTGTAATAATCTCT 568
Db 206043 ATGTGCTAACAGAAGATATATTTTATGTTTGTGAGTGAATGTTTGTGAGTATGT-T 205985
QY 569 AGGTCCACTTGGTTTATGACATCAGTGTAGCTCCAGCATTTCTCTGTTTGGTTTGT- 627
Db 205984 AGGTCCATTTGGTTTACATATTTGTTTACAGCATTTCTATGAAATCAGTTTAGTTT 205926
QY 628 ---GAGATGACCTAACTGTTGGAGAGAAAGGGTATTAAGTACGCCACTATCTGTGT 684
Db 205925 TTGTCTAATGATGCTATTGGCAATGTGTGGAATGAAGTCTCTATTATCAGGGTGT 205866
QY 685 GA-GTCAATATGATTTTACCTGCTAGCTGTGCTGTTTATGAAGTGGGTGACATTTG 743
Db 205865 GAGGTCAATATG---TTTAAAGCTGTAGTGTGTTTCTTTTAAAGACTGGGTG-CCTTA 205810
QY 744 TGTGTGTGCATAGACATTAAGAAATGCAATGTCTCTTGTGGATTTT---CCTTTGAT 800
Db 205809 TGCTTGGAGGATAGATGTTAAATAATGAATGTCTCTTGTGTTT---TTTCCCTTGAT 205750
QY 801 GCCTATGTAGTATCTTCCCAATCTCATCTGCTTAGTTTGGGTAAAGTCTA-TTAGTC 859
Db 205749 GAGTATGTAGTGTCTCTTATCTCTCTCTGATCAATTTTGAATTTGAAGTCTACTTTGTC 205690
QY 860 AGATATTAATAATGACTGTATCGGCTGCTCTTAGGGCAATTTGCTTAGAATATCTTTT- 918
Db 205689 AGAATTAATAATGACTGCACAGCTGCTCTTAGGTTCATTTGCTTGCATATCTTTT 205630
QY 919 ---CCATCCTTTTACTTAAGTGATGCTATCC---ATGGTGTGCTGTTTGGATG 973
Db 205629 TCCACATTTTACCTGAGGTGATGCTATCTCTAGATGTTAAAGTATATTTCTTAGATC 205570
QY 974 CAGCAGTAGGATGA-----TCTTGTGTTTCATATCCATCTGTTACCAG 1018
Db 205569 CAGCAGAAAGATGATCTTTTATGTTGTTTGTGTTTATTTTATTTCTGTTAACTG 205510
QY 1019 TATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGTTGAGAATTAATCAATGAGC 1078
Db 205509 CGTCTTTTATG-----GAGAAATGACCATTTATGTGCGAAGATTAATGATC 205459
QY 1079 AGTGTGTGATCTGTTATCTGTTATCTGACCTGTGGAAGTGTGTGTGTGTGTGTGT 1138
Db 205458 AGTGTGTGATCTGTTATCTGTTATCTGTTGTTGTTGTT-----TGTAGTAGTA 205415
QY 1139 GT 1198
Db 205414 TGGTGTGATGCTGGTGTGTGTGAACGTGGAGGTGGAGGTGTGTGTGTGTGTGTGTGT 205355
QY 1199 TCTGTGTTCTCTCCCTCTTTTGAATTTTGGCCTGGAATTAATTAATTAATTAATTTCT 1258
Db 205354 TGTGTGTCTGTTTCCAACTTTCACCTGCTGCTGTTGTTGTTTATTTCTTGTCTTTAT 205295
QY 1259 TGAATGTGGTGAACATCTTTAGATGAGTGTGTTTCTCTAGCCCTCTTTAGGTCTGCAIT 1318
Db 205294 TGGGTATGTTAAATGCTCTTTAGATGAGTGTGTTTCTCTAGCTCTGCTA---TAGAGTA 205238
QY 1319 TGAAGATAGATATTTCTTTACATCTGATTTTATCTAGAAATGCTTTCTTCTCCCAACTAT 1378
Db 205237 GATTCCTAGATATTTGCTTACATTTGGTTTATCATGTAATGCTTATTTTCTGTCTAT 205178
QY 1379 TGTGACAGAAAGTTTTTCTAAGTGCAGTGTCTGGCCTGACATCTGTAGTCTCTTGAAGT 1438
Db 205177 TGTGATTAAGGTTTTGCTGGGTATGATGATGCTAGGTGGCACCTATGATGCTCTAGAT 205118
QY 1439 CTGTAGACATCTGTGACAGGCCCTTCTTACATTTTGTAGTTTCTATTGGAAGGTCAGGTG 1498
Db 205117 CGGCAGACATCTGTCCAGGCCATCTGGCCCTTTAGAGTCTCCATTT-GAGNAGTTGGAAG 205059

QY 1499 TAATTTAAATACATCTGCCCTTATATGTTAAATGGTCTTTTTCCTTGCATCTTTTAAT 1558
Db 205058 AAATTTCTAAAATGCTGTCTTGACAGGTGACTTGGTCTTTTGGCTTGCATTTATA- 205000
QY 1559 ATTCTTTCTTTGTTCTATCTTTTAGTATTTAGATTTATTAATGACCTGTGGGAGTTTCTTT 1618
Db 204999 ---TTTTTCTTTGTTCTGTATTTAATGTTTGTAGATTTATGTGCGCAAGGAATTTCTT 204942
QY 1619 TTCGGGTCCAATCTATTTGGTGTGTTTGTATGCTTCTTGTACCTTTGATAGCATCTCTTTC 1678
Db 204941 TTCTGCTTAGTCTATTTGCTG---TATGTGCTTCTCTATCTTTAGAGTCTTCTCTTC 204885
QY 1679 TCAAGGTTAGGAATTTTCTTTTGGTTTCTTTGAAAATATTTTCCCTGCTTTTGACC 1738
Db 204884 ATTAGGTTAGGGAATTTTCTTTTATGCTTTTACTGTAATATTTTCTGTGGCTTGACC 204825
QY 1739 TG-----CCTTCTTCCCCTTCTCTATTTCCCTTTGGTTTTCCTAGTGTCTC 1785
Db 204824 TGTTCCTTCTCTCTTTTGTACCTATCATCTTTGATTTTATCTTTTCATAGTGTCCC 204765
QY 1786 TGGCTTCTCTGGATGTTTATGCTGCTGATTTATTTTAGACTTTAATCTTTCTTTGACCAAG 1845
Db 204764 AGATTTCTCTGGAATGTTTAAATGCCAGGAGT-TTTTAGATTTAACTTTTGTG-TCAGG 204707
QY 1846 TATCCATTTCTCTATCTTGTCTCTCTGCTGAGATTTCTCTCTCTATCTCTGTATTC 1905
Db 204706 TATCCATTTCTCTATTTGTCTCTCAATGCAITAGATTTCTCTTTTATCTCTGTATTT 204648
QY 1906 TGTCACTGAGGCTTGTCTGAGGTTCTCTG-TTGGGTTCTTAATTTTTCATTTCCAGAT 1964
Db 204647 TGTTAGTGAGGCTTGTCTCTGAGTTCTCTGTTAAGTTCCTATATATTTTTCATTTTAGAT 204588
QY 1965 TTTCTTCAAGTTGGGTTTGTATTAATTTCAATTTTCACTTTTCAAGTCTCTGAAATGTTT 2024
Db 204587 AGCTCTCAGTTGTGTTTCTTTTATGATTTGATTTTGAATTTTCAAGTCTCTGAGCTGTCT 204528
QY 2025 TACTCATTTTCTTCCAGTATTTACATTTTTCATAGTGTCTTAAATGGAATTTTTCATTT 2084
Db 204527 TATACA-TTTCTTCTTATGTTGTGTTTTCACAGATTTTCTTTAAGAGACTTATAAATTT 204469
QY 2085 CCTTTTCAAGGACCTTTTATGAATTCATAAAATGTATGTTAAAGTCTTTGCTTGTGCTT 2144
Db 204468 CCTTTTAAAGGACCTCTATCACACTCATAAAGCTATTTTAAAGTCAATATCTATGCT 204409
QY 2145 CAGCTATGTTGATTT-----CTCAGGCGCTATTTGAATAGGTTTTCAGAGGACA 2195
Db 204408 AAGCTATGTTTCTAGTATTTCCAGGACAGTTTGTGTAGATTTGCTGGGCTCTAGTGGGACC 204349
QY 2196 TATTGCTCTGTTGTTTATGTTGTGTTTCTTTGCTTTGGCATATAGAGGCTGAGTTTGGG 2255
Db 204348 TATTGCTCTGTTGTTTATGTTGTGTTTCTTCCACTGTTGTTTATGCTGATGCTAGGATCAGGTTTGAG 204289
QY 2256 ATGATTTGAATTTCTAGGTGCT 2276
Db 204288 AAGATTGCACCTCTAGGTGAT 204268

RESULT 35

AC115335

LOCUS

DEFINITION

AC115335

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC115335 243576 bp DNA linear HTG 13-NOV-2002
Rattus norvegicus clone CH230-202N2, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.
AC115335
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 243576)
Muzny, D., Marie, Metzker, M., Lee, A., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,

Anayalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dunn, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georegorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Guaratone, P., Healand, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Haviak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louisegh, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K., Nwackeleme, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plapper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Plopper, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, R., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wiczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, D., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

Unpublished
2 (bases 1 to 243576)
Worley, K.C.
Direct Submission
Submitted (17-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 243576)

REFERENCE AUTHORS TITLE JOURNAL

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 13, 2002 this sequence version replaced gi:22856763.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with 'Ns' to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRKC
Center clone name: CH230-202N2
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 226894 bases at least Q40
Consensus quality: 229443 bases at least Q30
Consensus quality: 231082 bases at least Q20
Estimated insert size: 231793; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 238369: contig of 238369 bp in length
* 238370 238469: gap of unknown length
* 238470 242092: contig of 3623 bp in length
* 242093 242192: gap of unknown length
* 242193 243576: contig of 1384 bp in length.
----- Location/Qualifiers
1. 243576
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-202N2"
232183..233645
/note="wgs_contig"
235011..236227
/note="wgs_contig"
237154..238369
/note="wgs_contig"

FEATURES

source
1. 243576
misc_feature
misc_feature
misc_feature

Query Match 30.7%; Score 698.901; DB 1; Length 243576;
Best Local Similarity 66.8%; Pred. No. 0.17;
Matches 1377; Conservative 0; Mismatches 561; Indels 123; Gaps 28;

QY	270	TTCCCTCTCTCTCCAAACACCTTCTATTTCCTGATTTCCTATCTTGCTCATTTTAACTC	329
Db	47891	TTCAATTCAGTCTAGAAACTCTTTATTTCTTTCTAGAAACTGAC-CCCTCTTATTC	47949
QY	330	AGTAGTACGTTGTTGGTTTCCATAAGTTGTAAGTTTCTGTTCTGTTGTTGTTG	389
Db	47950	AGTATTGAGCTTTCAGTTTCCATGAGTTTGTAACCTTTTCTCTGTTATCTGTTGCTG	48009
QY	390	TTGTTATCAGATTATTAAGCTGTGGTGCAGATAGACATAGATTATTTCAATTGTC	449
Db	48010	A---TATCCAGTTTATCAATTCATTTGGACAGATAGCTCCAGGGTGTATTTCATTTTC	48066
QY	450	TTTATCTGTCGAGACTTGTCTTTGTTTGAATAATGATTAATTTTGAGAGTTTCATA	509
Db	48067	TTGTGTCGTGTGAGCTTGTATCT-----TTGAAAAAGATTTCATG	48106
QY	510	GGGTGCTGCAGAAAGTACA-GCTTTTCTGTTTGGTGAATAGCTCTAATCTCT	568
Db	48107	ATGTGCTAAACAAGATATATATTTTAATGTTTCAGGTGAATGTTTGTGAGTAGT-T	48165
QY	569	AGGTCCACTTGGTTTATGATCAGTACAGTACCTCAGCATTTCTCTGTTTCTGTTT	627
Db	48166	AGGTCCATTTGTTTACATATTTGTTA-CTACAGCATTTCTTCAATTCAGTTAGTTT	48224

QY 628 ---GAGATGACCTAAGTCTTTGGAGAGATGGGTAATGAAGTAGCCCACTATCTGTGT 684
DB 48225 TTGCTAAATGATGTCATTTGGCAATTTGTGGAAATGAAGTCTCTCATTAACAGGGTGT 48284
QY 685 GA-GGTCAATATGATATTTAGCTGTAGCTGTGTGTGTATGAATCTGGGTGACATTTG 743
DB 48285 GAGGCTCAATATG---TTTAAAGCTGTAGTGTGTCTTTTAAAGAACTGGGGTGC-CTTA 48340
QY 744 TGTGTGGTGCATAGACATTAAGAAATGCAATGTCCTCTTTGGTGGATTT---CCCTTTGAT 800
DB 48341 TGCCTGGAGGATAGATGTTAAAAATGAAATGTCCTCTGTGTGTTTTTCCTCCCTTGAT 48400
QY 801 GCCTATGATGATATTTCTCCCAATCTCATCTGCTAGTTTGGGTAAAGTCTA-TTAGTC 859
DB 48401 GAGTATGATGATGTCCTTTTATCTCTCTCTCATCAAAATTTGATTTGAAGTCTACTTTGTC 48460
QY 860 AGATATTAATGATGATGATCGGCTGCTCTTAGGGCCATTTGCTTAGATATCTTTT- 918
DB 48461 AGAAATTAATGATGATGATGATGCTGCTCTTAGGGTCAATTTGCTTGCATATCTTTT 48520
QY 919 --CCATCCTTTTACTTAAGGTGATGCTATCC---ATGGTAGTGTGCTTTTTTGGATG 973
DB 48521 TCCCAATTTTACCCTGAGGTGATGCTATCCCTAGATGTTTAAAGTATATTTCTTAGATC 48580
QY 974 CAGCAGTAGGATGA-----TCTTGTCTTTCATATCCATCTGTTACCAG 1018
DB 48581 CAGCAGAAAGATGATCTTTTATTTGTTGTTTGTATTTTATTTATCTGTTTAACTG 48640
QY 1019 TATCTTTTCTAGAGAAATTAAGATCATTTAGTCAATGATGATGAGATTAATCAATGAGC 1078
DB 48641 CGTCTTTTATG-----GAGATGTGACCATTTATGTCGGAAGATTAATGATC 48691
QY 1079 AGTGTGTGATGATCTTTGATCTTTGATCTTTGATCTTTGATCTTTGATCTTTGATCT 1138
DB 48692 AGTGTGTGATGATCTTTGATCTTTGATCTTTGATCTTTGATCTTTGATCTTTGAT 48735
QY 1139 GT 1158
DB 48736 TGGT 48795
QY 1199 TCTGTGTGTCTCTCCCTCTTTTGTATTTTGGCTTGGATTAATTAATTAATTAATTTCT 1258
DB 48796 TGTGTGTCTGTCTTCCAACTTCACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 48855
QY 1259 TGAATGTGGTAAACATCTTAGATGAAGTTTCTCTAGCTTCTTCTAGCTTCTTCTGAT 1318
DB 48856 TGGGTATTTTAAATCTTTAGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 48912
QY 1319 TGAATAGATATCTTTACATCTGATTTTATCTTAGAATGCTTCTTCTTCTTCTCAACTAT 1378
DB 48913 GATTCCTAGATATGCTTACATTTGTTTATCAATGATGATGATGATGATGATGATGAT 48972
QY 1379 TGTGACGAAAGTTTCTTAAAGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1438
DB 48973 TGTGATGAAAGTTTGT 49032
QY 1439 CTGTAGACATCTGTGCGAGGCTTCTTACATTTGATTTTGTCTTCTTCTTCTTCTTCT 1498
DB 49033 CGGACGACATCTGTGCGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 49091
QY 1499 TAAATCTAATACATCTGCTTCTTATATGTAATGTTTCTTCTTCTTCTTCTTCTTCT 1558
DB 49092 AAAATCTAATATGCTGTGCTTGTGACAGGTGATGATGATGATGATGATGATGATGAT 49150
QY 1559 ATTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1618
DB 49151 --TTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 49208
QY 1619 TTTCCGTCCTAATCTATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1678
DB 49209 TTCTGTCTAGCTATTTGCTG---TATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 49265
QY 1679 TCAAGGTTAGGAAATTTTCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1738

DB 49266 ATTAGGTAGGAAATTTTCTTTTATCGTTTACTGTAAATATTTTCTGTGGCTTTGACC 49325
QY 1739 TG-----CCTTCTTCCCCTTCTCTATTCCTTTGCTTTGCTTTTGCATAGTGTCTC 1785
DB 49326 TGTTCCTTCTTCTTCTTGTACCTATCTTCTTGAATTTTATCTTTTTCATAGTGTCCC 49385
QY 1786 TGCCTTCTTCTTCTTGTGCTGCTGATTTATTTAGACTTAACATTTCTTTTGACCAAGG 1845
DB 49386 AGATTTCTTCTTCTTGTGCTGCTGATTTTAAAGCAGGAGT-TTTTAGATTTAACTTTT 49443
QY 1846 TATCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1905
DB 49444 TATCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 49502
QY 1906 TGTCACTGAGGCTTGTCTCTGAGGTTCTCTG-TTGGGTCTTAAATTTTCTTCAATTTCCAGAT 1964
DB 49503 TGTAGTGTGAGGCTTGTCTTCTGAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 49562
QY 1965 TTTCTTCAAGTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2024
DB 49563 AGCTCTCAGTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 49622
QY 2025 TACTCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2084
DB 49623 TATACA-TTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 49681
QY 2085 CTTCTTCAAGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2144
DB 49682 CTTCTTCAAGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 49741
QY 2145 CAGCTATGTTGCTT-----CTCAGGCTTATTTGTAAGGCTTCTTCTTCTTCTTCT 2195
DB 49742 AAGCTATGTTTCTGATTTTCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 49801
QY 2196 TATGTCTCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2255
DB 49802 TATGTCTCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 49861
QY 2256 ATGATTTGATTTCTAGGTGCT 2276
DB 49862 AAGATGCACTTCTAGGTGCT 49882

RESULT 36

AC129381/c

LOCUS

DEFINITION

Rattus norvegicus clone CH230-218C13, WORKING DRAFT SEQUENCE, 5

unordered pieces.

AC129381

AC129381.5 GI:30578753

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Rattus norvegicus

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 250030)

Munzy, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alsbrooks, S., Amir, A., Anguiano, D.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Devilla, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, I., Garza, M.,

Db 25246 TGCITGGAGGATAGATGTAAATAATGAAATGCTCTCTGTTTTTTTTTCCCTTGAT 25187
QY 801 GCCTATGATGATATCTCCCAATCTCACTGCTAGTTTTGGGTTTAAAGTCTA-TTAGTC 859
Db 25186 GAGTATGATGCTCTCTTAATCTCTCTGATCAATTTGATTTGAAGTCTACTTTGTC 25127
QY 860 AGATATTAATAATGACTGATCGGCTTGCTTCTTAGGGCCATTTGCTTAGAATATCTTTT- 918
Db 25126 AGAAATTAATAATGACTGCACGAGCTTGCTCTTAGGTTCA-TTCTTGCAATATCTTTT 25067
QY 919 --CCATCTTTTACTCTAAGGTGATGCTATCC---ATGGTAGGTGCTTTTTTGGATG 973
Db 25066 TCCCACTTTTACCTGAGGTGATGCTATCTCTAGATGTTAAAGTATATTTCTTAGATC 25007
QY 974 CAGCAGTAGGATGGA-----TCTGTGTTTTTATATCAATCTGTTACCCAG 1018
Db 25006 CAGCAGTAAGATGATCTTTTATTTGTTGTTTGAATTTTATTTCTGTTAACTG 24947
QY 1019 TATCTTTTCTAGAGAAATTAAGATCAATGATGATGATTTGAGAAATATCAATGAGC 1078
Db 24946 CGTCTTTTATG-----GAGAAATGTGACCAATTTATGTCGGAAGATATTAATGATC 24896
QY 1079 AGTGTTCTGGATCTCTGTATCTGACCTTGTGAAGTGTGTGTGTGTGTGTGTGT 1138
Db 24895 AGTGTTCTGATTTGCTATTTTGTGTTGT-----TGATGATGGA 24852
QY 1139 GT 1198
Db 24851 TGGT 24792
QY 1199 TCTGTGTGTCTCCCTCTTTTGTATTTTGGCTGGAAATTTATTTATTTATTTATTTCT 1258
Db 24791 TGT 24732
QY 1259 TGAATGTGGGTAACATCTTTAGATGAAAGTTTTTCTCTAGCCCTTTCTTAGGCTGTCAT 1318
Db 24731 TGGGTATGTTTAAATGCTTTAGATGAGTTTTCTCTAGCTCTGCTA---TAGAGTA 24675
QY 1319 TGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATGCTTTCTTCTCCAACTAT 1378
Db 24674 GATTCCTAGATATGCTTACATTTGGTTTATATCTAGATGATCTTTCTGCTCTAT 24615
QY 1379 TGTGACAGAAAGTTTTTCTAGTGCAGTAGTCTGCGCTGACATCTGTAGTCTCTTGAGT 1438
Db 24614 TGTGATGAGGTTTTGCTGGGTATGATGATGCTAGGCTGGACCTATGATGCCITAGAT 24555
QY 1439 CTGTAGACATCTGTCAGGGCCCTTCTTACATTTTGTAGTTTCTATTTGAAAAGTCAGGTG 1498
Db 24554 CGGCAGCACATCTGCCAGGCCATTTCTGGCCCTTTAGAGTCTCCATTT-GAGAAGTTGGAAG 24496
QY 1499 TAAATCTTAATACATCTGCTTTATATGTAATGCTTTTCTTTTCTGCTGCTATCTTTAAT 1558
Db 24495 AAAATCTTAAATGCTGTCTGTCGACAGGTGACTTGTCTTTTGTGCTTGCACTTTTATA- 24437
QY 1559 ATCTCTTCTTCTTCTATACITTTTATAGTGAATTTGATTTATATGCACTGTTGGGAGTTCTT 1618
Db 24436 -TTTTCTTGTCTGTATATTTAAATGCTTTAGATTATATGTCGCAAAAGGATTTCTT 24379
QY 1619 TTCGGTCCAAATCTATTTGTTGTTTGTATGCTTCTTGTACCTTGTAGAGGATCTCTTTC 1678
Db 24378 TCTGCTCTAGCTATTTTGTCTG- --TATGTGCTTCTTATATCTTTTAGAGTCTTCTCTCTC 24322
QY 1679 TCAAGTTTAGGAAATTTTTCTTTTGTGTTTCTTGTGAAATTTTCTCTCTCTTTTGACC 1738
Db 24321 ATTAGTTAGGAAATTTTCTTTTATCTTTTACTGTGAAATTTTCTGTGGCTTTGACC 24262
QY 1739 TG-----CCTTCTCCCTTCTCTATCTCTTTGTTTGTGATGCTCTC 1785
Db 24261 TGTCTTCTCTTCTCTACTTTGTACCTATCATCTTTGATTTTATCTTTTCTATGTTGCTCC 24202
QY 1786 TGGCTTCTGATGTTTTATGCTGATATTTTAGACTTAACATTTTCTTTGACCAGG 1845

Db 24201 AGATTTCTGGAATTTTAAATGCCAGGAGT-TTTTAGATTTAACTTTTTTTTGTG-TCTAGG 24144
QY 1846 TATCCATTTCTTCTATCTTCTGCTTCTACTGCTCAGATTTCTCTCTCTATCTCTCTGATTC 1905
Db 24143 TATCCATTTCT-TCTATGTCCTCCAAATGCAATGATTTCTCTCTTTTATCTCTGATTT 24085
QY 1906 TGTCAAGGCTTGTCTCTGAGGTTCTCTG-TTGGGTTCTTAAATTTTCTCAATTCAGAT 1964
Db 24084 TGTTAGTAGGCTTGTCTCTGAGGTTCTCTGAGGTTCTCTGAGGTTCTCTGAGGTTCTCTGAG 24025
QY 1965 TTCTCTCAAGTTTGGGTTTGTGTTTAAATTTTAAATTTTCCATTTTCCAGTCTCTGAAATGTTT 2024
Db 24024 AGCTCTCAAGTTTGTGTTTCTTTTATTTGATTTTGTATTTTCAAGTCTCTGAGTGTCT 23965
QY 2025 TATCTCATTTTCTCCCAAGTATTTTACATTTTCTCATAGGTTTCTTTAAATGGAATTTTATTCATTT 2084
Db 23964 TATACA-TTCTCTTCTATTTGTTGTTTTCACAGATTTCTTTTAAGAGACTTATAAATTT 23906
QY 2085 CTTCTTCAAGGACTTTTATGAATTCATTAATATGATATGATTAAGTCTCTGCTTGTGCTT 2144
Db 23905 CTTCTTTAAGGACTCTATCACACTCAATAAAGCTATTTTAAAGGCTATTTCTCTATGCT 23846
QY 2145 CAGCTATGTTGCAAT-----CTCAGGGCTATTTGTAATAGGTTTCTTAGCAGGACA 2195
Db 23845 AAGCTATGTTTCAATTTCCAGGACAGTTGTGTAGAGTTGCTGGCTCTAGTGGGACC 23786
QY 2196 TATGTCCTGTTTGTATTTGCTGTTTGTGTTTGTGTTTGGCATATAGACGCTGATTTGG 2255
Db 23785 TATGTCCTGTTGCTATCCACTGTTGTTTATGCTGATGCTCTAGCATCAGGTTTGTAG 23726
QY 2256 ATGATTTGATTTCTAGGTTCT 2276
Db 23725 AAGATTGCATTTCTAGGAT 23705

RESULT 37

AC125973 228302 bp DNA linear HTG 13-MAY-2003
LOCUS Rattus norvegicus clone CH230-64N5, *** SEQUENCING IN PROGRESS ***,
DEFINITION 5 unordered pieces.
AC125973
VERSION AC125973.3 GI:30579661
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Bukaryota; Metazoa; Chordata; Vertebrata; Eutelestomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 228302)
AUTHORS Muzny, D., Marle, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Ayalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandar-Araki, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, B., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, F., Geer, K., Gill, R., Grady, W., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorensheva, L., Loulsegged, H., Lozano, R.J., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
 Nwokilemen, O., Okwou, G., Olarunpunsagoon, A., Pal, S., Pank, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puato, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sibson, I., Sitter, C.D., Smales, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wiczyski, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GLNF
 Center clone name: CH230-64N5
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 210267 bases at least Q40
 Consensus quality: 214022 bases at least Q30
 Consensus quality: 217194 bases at least Q20
 Estimated insert size: 221667; sum-of-coverage
 Quality coverage: 6x in Q20 bases; sum-of-coverage estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 215461: contig of 215461 bp in length
 * 215561: gap of unknown length
 * 215562: contig of 2452 bp in length
 * 218013: gap of unknown length
 * 218014: contig of 1263 bp in length
 * 219376: gap of unknown length
 * 219377: contig of 1870 bp in length
 * 221347: gap of unknown length
 * 221348: contig of 6856 bp in length.
 * 221447 228302: contig of 6856 bp in length.
 * 221447 228302: Location/Qualifiers
 source
 1. 228302
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-64N5"
 Query Match 30.6%; Score 697.602; DB 1; Length 228302;
 Best Local Similarity 66.1%; Pred. No. 0.17;
 Matches 1282; Conservative 0; Mismatches 554; Indels 104; Gaps 20;
 QY 270 TTCCCTCTCTCTCCAAACACCTTCTATTTCTTGAATTTCTATCTTGGCTCATTTTAACTC 329
 |||||
 DB 84182 TTCCCTGAATCTAGAAAGTTTAAATTTCTTTATTTTCTTCCCTGACCCAGAGATCATG 84241
 QY 330 AGTAGTGTGTTGTTGTTTCCATTAAGTTTGAAGTTTCTGTTGTTCTGTTGTTGTTG 389
 |||||
 DB 84242 GATAAAGAGTTGTTCTATTTCCCTGAGTTTGAAGCTTTTGGTGTATGTTGTTGAAG 84301
 |||||
 QY 390 TTGTTATCTAGATTTAAAGCTGTGGTGGTCAGATAGGACATAGATATTTCAATTTGTC 449
 |||||
 DB 84302 -----TCCAGTCTTAATCCATGGTGGTCTGATGATACAGGATTTTCAATTTTC 84355
 QY 450 TTTTATCTGTCGAGACTTCTTGTGTTTGAATATGATTTCAATTTTGGAGA--GTTTCA 507
 |||||
 DB 84356 TTGTTATCTGTTGAGGCTTGTGTTTGGCCCAAGTATATGTTTGGAGAACGTTTCA 84415
 QY 508 TAGGGTCTGACAGAGCTGATGCTGTTGTTGTTGGTGAATAGTCTGTAATATCTC 567
 |||||
 DB 84416 CAAGGTGCTGAGAGAGGATGATTTCTTTTGTATTTG-----GGTGAATTTG 84463
 QY 568 TAGGTCCACTTGGTTTATGACATCAGTTAGCTCCAGCAATTTCTGTTTCGTTTTGTT 627
 |||||
 DB 84464 TTGGGTCCTTTGATTCTATAAATCAGTTAGTTCATTATTTCTGCTTAGTTTTGTCT 84523
 QY 628 GAGATGACCTAATGTTGAGAGAAATGGGATTTGAGTAGCCCACTATCTGTTGAG 687
 |||||
 DB 84524 GGATAACCTGTTGAGTTGGTGGAGAGTAGGTTGTTGAAGTCTCTCATTATTAATGTCGG 84583
 |||||
 QY 688 G-TCAATATGCTATTTTAGCTGTAGCTGTGCTTTGTTTATGAACCTTGGGTGACATTTGT 746
 |||||
 DB 84584 GTTTGACGTGTGATTTATGCTTTAGCAATGTTCTTTAACAATATGTTGGTGGCTTGCAT 84643
 QY 747 TTGGTGATAGACATTAAGAAATGCAATGCTCTTGGTGA--TTTTCTTTGATGCCTA 805
 |||||
 DB 84644 TTATGGCTAGATGTTTCAAGAACTGAGATAAATCTCGGTATATTTTCTTTGATGACTA 84703
 QY 806 TGTAGTATTTCTTCCCAATCTCATCTCTTAGTTTGGGTTTAAAGTCTA--TTAGTCAGATA 864
 |||||
 DB 84704 TGAAGTGTCTTTGCTGCTCTTTTGAATTTGTTTAAAGTCTATTTTATTAGATA 84763
 QY 865 TTAATATGATGCTGATCGGCTTGTCTTCTTAGGGCCATTTGCTTAGAATA--TCTTTTCCATC 923
 |||||
 DB 84764 TTAGCATGGTACTTTCAGCTTGTCTTGGGACCAATTTGTTGGAAAACTTTTCTTCTACC 84823
 |||||
 QY 924 CTTTACTTCTAGGTGATGCTATC--CATGGTAGTTGTTCTTTTTCGATGCAGCACT 980
 |||||
 DB 84824 CCTTTACTCTAAGGCAATGCTATCTTTCTGCTAGGTTGTTCTTAGTATGCAGTAGA 84883

981 AGGATGGATCTTGTGTTTTCATATCCATCTGTTTACCCAGTATCTTTTCTAGAGAAATTA 1040
84884 ATGTTGGGGTCTGTTTTCACAGTATCTATTAATCTGTGTCCTTTTATTTGGGGATTTGA 84943
1041 GATCATGTAGTCATTTGAGTTGAGAAATTAATCAATGAGCAGTGTGTTGGATCTTGTAT 1100
84944 G-----TCCATTTGATCTAGAGATATTAATGACCATGATTTGCTAAATTTCTGATAT 84995
1101 CTTGCACTTTGCAAGT 1160
84996 TTTGATGTTGGT 85055
1161 GT 1220
85056 GTGTG-----CACGGCAGCGCTCTTCCCTGTTTGTG 85088
1221 GATTTTGGCCTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1280
85089 CTCCTGCTGGTAAATCTTATTTCCGTGTTTCTTGGATGTTGTTTACTCCTTGG 85148
1281 ATTGAAGTTTTCCTCTAG--CCTTCTTTAGGTCTGCATTTTGAAGATAGATATCTTTAC 1338
85149 GTTGAGTTTTCCTTTGAGTATTTCTGTGAAGCTGGATTTGTGATAGATATCTTTAA 85208
1339 ATCTGATTTTATCTTAGAATGCTTCTTTCTTCCAACTATTTGTGACAGAAAGTTTCTTA 1398
85209 ATTGGTTTGTCTTGAATTAATGTTTCTTCTATCTATGATGATGATGAGAGTTTTCG-- 85266
1399 AGTCAGATAGTCTGGCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGG 1458
85267 --TCTAGTAGTCTATGTTGAGATCTGTGATGTTTGTAGAGCTACAGACATCTGTTCAGG 85324
1459 GCCTCTTACATTTTGTAGTTTCTATTTGGAAGAGTCAGTGTAAATCTTAAATCAATCTGCT 1518
85325 CCCTCTTAGATTTTGTAGTCTCTGTTGG--GAAGTTGGGTGTAATCTGATTAATCTGCT 85383
1519 TTATATGTTAAATGCTTTTTCCTTCGTCATCTTTTAAATTAATTAATTAATTAATTAAT 1578
85384 TTGTATGTT-AGTGGGCCCTTCTTCTGGAAGCTTTTAAATTAATTAATTAATTAATTAAT 85442
1579 TTTTGTAGTATTTGATTATTAATGATCTGTGGGAGTTTCTTTTCCGGTCCAATCTATTGG 1638
85443 ATTTACTGTTTGTAGATTTTGTGA--GCTGAAGGATTTCTTTCTCATCTAATTAATGG 85501
1639 TGTTTGTAGTCTCTTG-----TACTTGTAGTAGGATCTCTTTCTCAAGTTAGAA 1691
85502 TGTCTGTAGCTCTCTGTATGTTTACTTTTACAGGATCTCTTCTTTAGGTTTGTAA 85561
1692 ATTTTCTTTTGTGTTTCTTGAAATATTTTCCCTGCTTTTGACCT-GCCTTCTTCCC 1750
85562 TGTTTCTTCTATGATCTTGTGAAATATTTTCTAGTCTTGGAACTGGCACTCTTTAG 85621
1751 CTTCTCTATCTCTTTGTTTGTGATAGTGTCTCT-----GGC 1789
85622 CTTCTCTATCTCTATTTATTTCTTAGTCTAGTCTTTTATAGTATCTCTAGATGATAGTA 85681
1790 TTCTCGGATTTTATGCTGCTGATATTTTATGACTTACATTTTCTTTGACCAAGGTATC 1849
85682 TCTTGGATGTTTGTGTGAGAAATCTTTTATGATTTAAATTTTCTTGTACGATGATATC 85741
1850 CATTTCTCTATCTTGTCTTCACTGCTGAGATTTCTCTTCTATCTCTTGTATTTCTGTC 1909
85742 AATTTCTCTATAGCATCTCTAGACCTGAGATTTCTTCTATCTCTCTGTGTTCTGTT 85801
1910 AGTAGGCTTGTCTCTAGAGTTCTGTGGGTTCTTAATTTTCTTCAATTCAGATTTCTCT 1969
85802 GGTGATGCTTGTATCTGTTTGTCCCGTCTCTCCCTAGTTTCTATCTCTAGCATCTCC- 85860
1970 TCAGTTGGGTTTGTATTAATTAATTTCTTCCACTTTTCAGTCTGCTGAAATCTTTTACTC 2029
85861 -CAGTTTGTGTTTCTTATGCTTCTATTTCAATTTTGTAGTCTGTGACAGCTTTATTT 85919
2030 ATTTTCTCTCCAGTATTTACATTTTCAATAGGTTTCTTTAAAGGA--TTTATTCATTTCTCT 2087

Db 85920 ATTTCTTTCACCTGTTGTTTATTTTCTCTGATTTTCTTTAAGGATTTTATTTGTTTCTCT 85979
Qy 2088 CTTCAAGGACCTTTATGAATTTCAAAAATGATGTTAAGTCTCTGCTTGTGCTTCTCAG 2147
Db 85980 CTTTAAATGCTCTACCTGTTTGGTAGTATTTTCTCTGATTTTCTTTTAAAAAAGGTTTAT 86039
Qy 2148 CTATGTTGCACTTCTCAGGCG 2167
Db 86040 TCATTCTCTTTAAGGCC 86059

RESULT 38
AC134939

LOCUS AC134939 183078 bp DNA linear HTG 15-NOV-2002
DEFINITION Rattus norvegicus clone CH230-375A21, WORKING DRAFT SEQUENCE, 3
unordered pieces.

AC134939 GI:25007201

VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.

KEYWORDS Rattus norvegicus (Norway rat)

SOURCE

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutharia; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 183078)

REFERENCE

AUTHORS

Muzny, D., Marie, M., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Altschrocks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhat, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, A.,
Chacko, J., Chavez, D., Chen, R., Chen, Y., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M., Davis, C., Davy-Cartroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, S.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, T., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S., L., Hodgson, A., Hoques, M.,
Hollins, B., Howell, S., Hulyk, S., Hume, J., Idrebird, D., Jackson, A.,
Jackson, L., Jacob, B., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovac, C.,
Kowis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshewa, L., Loulsegad, H., Lozado, E., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martin, E.,
Mawhinney, S., McLeod, M., McNeill, T., Z., Neenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nuckeleneh, O., Okwuonu, G., Olarnpungsoo, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Prannkoc, C.,
Plopper, F., Poldexter, A., Popovic, D., Primus, E., Pu, L., L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, R., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von

Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
Weinstock, G. and Gibbs, R.A.
Unpublished
2 (bases 1 to 183078)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (03-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 183078)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23477788.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

REFERENCE AUTHORS TITLE JOURNAL

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: KCMW
Center clone name: CH230-375A21
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 168021 bases at least Q40
Consensus quality: 170281 bases at least Q30
Consensus quality: 171978 bases at least Q20
Estimated insert size: 170392; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 180012: contig of 180012 bp in length
* 180013 180112: gap of unknown length
* 180113 181340: contig of 1228 bp in length
* 181341 181440: gap of unknown length
* 181441 183078: contig of 1638 bp in length.

FEATURES source

1. 183078
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-375A21"

misc_feature

1. 1242
/note="wgs end_extension
clone_end:T7"

misc_feature

5362..6656
/note="wgs end_extension
clone_end:T7"

misc_feature

6707..7817

		/note="wgs end_extension clone_end:T7" complement(8960..9793) /note="clone_boundary clone_end:T7" site: end sequence:RXAW111TJ" 26621..173267 /note="clone_boundary clone_end:Sp5 site: end sequence:BZ263204" 173661..175381 /note="wgs end_extension clone_end:Sp6" 178477..180012 /note="wgs end_extension clone_end:Sp6"
	Query Match Best Local Similarity 66.2%; Pred. No. 0.15; Matches 1371; Conservative 0; Mismatches 585; Indels 115; Gaps 27;	
Qy	270 TTCCCTCTCTTCCAAACACCTCTATTCTTCTGATTTCTATCTTGGCTCATTTTAACTC 329	
Db	42317 TTCAATGAATCTAGAAAATCTTTAAAGTTCTTTTATCTGTCTTGATTCATTTTCATTC 42376	
Qy	330 AGTAGTCAGTTGTTTGGTTTCCATAAGTTTGTAAAGTTTCTGTGTTTCTGTGTTGTTG 389	
Db	42377 AATAGAGAGTTGTTTACTTTTATGAGTTTGAAGATTCTAATGTTCTGTCATTTG--- 42433	
Qy	390 TTGTTATCTAGATTAAAGCTGTGGTCAGATAGACATAGATGATTAATTTCAATTGTC 449	
Db	42434 -TGAATCCAGCTTTAATCCATGATGATTAACAGGATGACAGGAGTTTCTATTTTC 42492	
Qy	450 TTTTATCTGTCGAGACTTGTCTTTGTTTGAATAATGATTAATCAATTTGGAGAGTTTCATA 509	
Db	42493 TTATATCTGTTAAACATTTGCTTTCTCTGAGTATGCTCAATTTTAGGAGNAGTCTTA 42552	
Qy	510 GGGTGTCTGAC--AAGAAGGTACAGTCTTT--GFGTTTTGGTGAATAGTCTGTAATATC 565	
Db	42553 TGAGGCATACAGAAAGAGTTTATTCTTTTGTGTTGAATGAATGTAATGTTCTGTAATAATC 42612	
Qy	566 TC-TAGTCCACTTGGTTTATGACATCAGTTAGTCCAGCATTTCTCTCTGTTGGTTTTT 624	
Db	42613 TCTGGGTCCATTGGTTTGTATTAACATCTGTTAGTCCATATTTCTCTGTTAGATTTTA 42672	
Qy	625 GTTGAGATGACCTAACTGTTGGAGAGATGGGGTATTGAAGTAGCCCACTATCTGTGTG 684	
Db	42673 TCTG-TATGACCTGCCCTATTGATGAGATGGGGTAAATGAAGTCTCCCATTAATCATTTGT 42731	
Qy	685 -GAGGTCAATATGATTTTAGCTGTAGCTGTGCTTTTATGAACTTTGGGTGGGTGACATG 743	
Db	42732 GGGGGTCAATATGATTTTAAAGCTATAGTAGTAATTTCTATTACAAATTTGGGFGCTCTTG 42791	
Qy	744 TGTTTGTGATAGACATTAAGAAATTCGAATGCTCTTCTGGTGA--TTTTCCTTTGATGC 802	
Db	42792 TGCTTGGCGGCAAAATTAAGAAATTCGAATTCCTTCTTGGTTGAGTTTCTCTGGATCA 42851	
Qy	803 CTATGATGATTTCTTCCCAATCTCATCTGCTGTAGTTTGGGTTTAAAGTCTATTAGTC-AG 861	
Db	42852 ATATGGGGTGTCTTCCCTATCTCTTTTGATAATTTTGGTTTGTAGAGATTTTCACTAG 42911	
Qy	862 ATATTAATGACTGTATGGCTTGTCTTGGGGCAATTTGCTTAGAATATCTTTTCCA 921	
Db	42912 ATATTAATGCTATAGATGACCTTACTTCTAGTCCATTTGCTTGGAAATATCTTTCCA 42971	
Qy	922 TCCTTTTACTTAAAGTGATGCTATCCATGGTAGGTTGTCTTTTGGATGACGAGTA 981	
Db	42972 ACCCTTTAC-CTTGGGGTAATGTCTGTCTTGATGTTTCAGATTGTCTCTTGTCTACAGTA 43030	
Qy	982 GGATGGATCTGTTTTCATATCCATCTGTTACCCAGATCTTTTCTTAGAGAAATTAAG 1041	

Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Usmani, K., Valas, R., Verà, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission

Unpublished

2 (bases 1 to 199545)

Worley, K.C.

Direct Submission

Submitted (21-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 199545)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 19, 2002 this sequence version replaced gi:23809119.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GVMU

Center clone name: CH230-447P6

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 180908 bases at least Q40

Consensus quality: 182758 bases at least Q30

Consensus quality: 183743 bases at least Q20

Estimated insert size: 184876; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html). NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 198343: contig of 198343 bp in length

* 198344 198443: gap of unknown length

* 198444 199545: contig of 1102 bp in length.

Location/Qualifiers

1. .199545

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-447P6"

1. .1634

/note="wgs end_extension

clone_end:T7"

misc_feature

misc_feature

14358. .15643

/note="wgs end_extension

clone_end:T7"

17817. .18653

/note="clone_boundary

clone_end:T7

site:

end_sequence:BZ250252"

complement(197253. .197758)

/note="clone_boundary

clone_end:Sp6

site:

end_sequence:BZ250253"

Query Match

30.4%; Score 692.504; DB 1; Length 199545;

Best Local Similarity 66.2%; Pred No. 0.16;

Matches 1371; Conservative 0; Mismatches 585; Indels 115; Gaps 27;

Qy 270 TTCCCTCTCTCTCCAAACACTTCTATTCTTCTGATTTCTATCTTGCTCAATTTTAACTC 329

Db 102775 TTCATTCAATCTAGAAATCTTTAAAGTCTTTTATTCTGCTGTGATTCATTTTCATTC 102716

Qy 330 AGTAGTCAGTGTGTTGGTTTCCATAGTTTGTAAAGTTTCTGTTGTTCTGTTGTTGTTG 389

Db 102715 AATAGAGAGTGTGTTTACTTTTTTATGAGTTTGTAAAGATTTCTAAAGTTTCTGTCAT 102659

Qy 390 TTGTATTATAGATTTTAAAGCTGTGCTGTCAGATAGGACATAGAGATTTATTTCAATTGTC 449

Db 102658 -TGAATCCAGCTTTAATCCATGATGATTAACAGGATGCGAGGAAGTTATTTCTATTTTC 102600

Qy 450 TTTTATCTGTCGAGACTTGCTTTGTTTGAATATGATTTCAATTTTGGAGATTTTCATA 509

Db 102599 TTATATCTGTAAACACTTGCTTTCTGCTGAGTATGCTGTCATTTTAGAGGAAGCTCTA 102540

Qy 510 GGGTGTCTGAC - AAGAGGTACAGTCTTT - GTGTTTTGTGTAAGTAGCTGTGTAATAATC 565

Db 102539 TGAGGCATACAGAGAGAGGTTTATTTCTTTGTTGTTGAATGAATGTTCTGTAATAATC 102480

Qy 566 TC-TAGTCCACTTGGTTTATGACATCAGTTAGCTCCAGCATTTCTGTTGTTGTTT 624

Db 102479 TCTTGGTCCATTTGGTTTATAACATCTGTTAGTTCCTGTTGTTGTTGTTGTTGTTGTTG 102420

Qy 625 GTTGAGATGACCTAACTGTTGGAGAGAAATGGGGTATTGAAGTAGCCACATCTATCTGTGTG 684

Db 102419 TCTG-TATGACCTGCCCTATTGATGAGATTTGGTAAAGAGTCTCCCATATCATTTGTTG 102361

Qy 685 -GAGGTCAATATGATGATTTTATGCTGTGCTGTTTATGAACTTGGGTGACATTTG 743

Db 102360 GGGGGTCAATGTGATTTAAAGCTATAGTAGTATTTCTATTACAAATTTGGGTGCTCTTG 102301

Qy 744 TGTTTGTGTGATAGACATTAAGAAATTCATGTCCTCTCTGTTGGA-TTTTCTCTTTGATGC 802

Db 102300 TGCTTCGGCCGAAATATTAAAGAAATTTGAAATTCCTTCTTGGTTGAGTTTTCTCTTGGATGA 102241

Qy 803 CTATGTAGTATTTCTCCCAATCTCATCTGCTTATGTTTGGGTTTAAAGTCTATTAGTCT-AG 861

Db 102240 ATATGGGGTCTTCTCCCTATCTCTTTTGATAATTTTGGTTTGTAGACTATTTCCTAG 102181

Qy 862 ATATTAAATAGACTGTATCGGCTTCTTATAGGCAATTTGCTTAGAAATATCTTTTCCA 921

Db 102180 ATATTAAATGTCTACATCATGCTTACTTCTTAGTCCATTTGCTTGGAAATATCTTTTCCA 102121

Qy 922 TCTTTTACTCTAAGGTGATGCTATCCATGTTAGTGTGTTTGTGTTTGGATGAGCAGTA 981

Db 102120 ACCCTTTAC-CTTGGGGTAAATGCTGCTCCCTGATGTTTCAGATTTGCTGTTGTCAGTA 102062

Qy 982 GGATGGATCTGTTTTCATATCCATTTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAG 1041

Db 102061 GGAGGGTAGATGCTGT-TTTTCAATTTCTGTAGACCATGCTCTTTTATTGGGGAATTGAG 102003

Qy 1042 ATCATTGAGTCAATGATGTTGAGAATATCAATGACAGCTGTTTGGGATCTCTGTTATC 1101

Db 102002 ACC-----ATATTGAGAGATATCGATGACACTGATTGTTGATTCTTGTATT 101955

FEATURES

source

[illegible]

QY	2138	TGTGCTTCAGCTATGTTGGCA-----TTCTCAGGCGCTATTGTAATAGGGTTTTTACG	2188
DB	100903	TCTACTTTGGTTGTGTTGGAAATATTCTGGGTCGTCTGGTGGAATAGATGGCTCTACT	100850
QY	2189	AGGGACATATTGCTCGTGTGTTATGTCGTGTTTTTGCCTTTGGCATATACAGGCTGA	2248
DB	100849	CGAGACATATTGCTCGTGTGTTTCTTGATCATGTTTTTATGCTTCCTTTGGAATCTGGA	100790
QY	2249	GTTTGGGATGATTGTAATCTTAGGTGCTGAT	2279
DB	100789	ATTAAGATGATTTAAAAGCTAGGTGCTGGT	100759

RESULT 40	ACI29417/c
LOCUS	Rattus norvegicus clone CH230-287M3, WORKING DRAFT SEQUENCE, 6
DEFINITION	unordered pieces.
ACCESSION	ACI29417
VERSION	ACI29417_4 GI:25082889
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
	Rattus.
REFERENCE	1 (bases 1 to 213532) Abramzon,S., Adams,C., Alder,J., Muzny,D.,Marie,, Metzker,M.Lee., Allen,A., Anguitano,D., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguitano,D., Ayalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blanchburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dedrich,B., Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Faves,K., Egan,A., Escoto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Georegeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Huliyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovac,C., Kowts,C., Kratt,C.L., Lebow,H., Levant,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Loulsegad,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munica,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwakilemech,O., Okwuonu,G., Olarnpusagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., L., Fuazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,K., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Sosa,J., Steinkle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,I., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villaseña,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,

Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Ton, K., Liu, G., Lu, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (03-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 3, 2003 this sequence version replaced gi:20219162.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIER

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L20306
Center clone name: 442.M.5

----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 186305 bases at least Q40
Consensus quality: 188311 bases at least Q30
Consensus quality: 188966 bases at least Q20
Insert size: 194000; agarose-fp
Insert size: 189347; sum-of-contigs
Quality coverage: 6.4 in Q20 bases; agarose-fp
Quality coverage: 6.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 313: contig of 313 bp in length
* 314 413: gap of 100 bp
* 414 1100: contig of 687 bp in length
* 1201 1200: gap of 100 bp
* 1201 2307: contig of 1107 bp in length
* 2308 2407: gap of 100 bp
* 2408 3409: contig of 1002 bp in length
* 3410 3509: gap of 100 bp
* 3510 5052: contig of 1543 bp in length
* 5053 5152: gap of 100 bp
* 5153 72089: contig of 66937 bp in length
* 72090 72189: gap of 100 bp
* 72190 74638: contig of 2449 bp in length
* 74639 74738: gap of 100 bp
* 74739 77109: contig of 2371 bp in length
* 77110 81155: contig of 3946 bp in length
* 81156 83255: gap of 100 bp
* 83256 90890: contig of 7635 bp in length
* 90891 90991: gap of 100 bp
* 90991 103547: contig of 12557 bp in length
* 103548 103647: gap of 100 bp
* 103648 120326: contig of 16679 bp in length
* 120327 120426: gap of 100 bp
* 120427 132054: contig of 11628 bp in length
* 132055 132154: gap of 100 bp

```

```

* 132155 148505: contig of 16351 bp in length
* 148506 148605: gap of 100 bp
* 148606 169417: contig of 20812 bp in length
* 169418 169517: gap of 100 bp
* 169518 189830: contig of 20313 bp in length
* 189831 189930: gap of 100 bp
* 189931 190947: contig of 1017 bp in length.

```

FEATURES
Location/Qualifiers
source
1. 190947

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-442M5"
/clone_lib="RPCI-23 Female Mouse BAC"
1. 313
/note="assembly_fragment"
clone_end:SP6
vector_side:left"

misc_feature

misc_feature 414..1100
/note="assembly_fragment"
1201..2307
/note="assembly_fragment"
2408..3409
/note="assembly_fragment"
3510..5052
/note="assembly_fragment"
5153..72089
/note="assembly_fragment"
72190..74638
/note="assembly_fragment"
74739..77109
/note="assembly_fragment"
77210..83155
/note="assembly_fragment"
83256..90890
/note="assembly_fragment"
90991..103547
/note="assembly_fragment"
103648..120326
/note="assembly_fragment"
120427..132054
/note="assembly_fragment"
132155..148505
/note="assembly_fragment"
148606..169417
/note="assembly_fragment"
169518..189830
/note="assembly_fragment"
189931..190947
/note="assembly_fragment"
clone_end:T7
vector_side:right"

Query Match 30.2%; Score 688.903; DB 1; Length 190947;
Best Local Similarity 65.5%; Pred. No. 0.16;
Matches 1343; Conservative 0; Mismatches 617; Indels 89; Gaps 24;

```

QY 270 TTCCCTCTCTCTCCAAACACCTTCTATTCTTCTGATTTCTATCTTCTGCTCATTTTAACTC 329
Db 59214 TTCAATTGAATTATAGAAATTCCTTAATTTCTTATTATTCTTCTGACCTAGTGGTCTTG 59273
QY 330 AGTAGTGAGTCTTTTGGTTTCCATAAGTTTGTAAGTTTCTGTTTCTGTTTGTGTTG 389
Db 59274 AGTAGGAGTCTTCTCAGTTTCCATGATTTCTGATGATTTCTGTTTGTGTTTGTGTTT 59329
QY 390 TTGTATCTTAGATTTAAAGCTGTGGTGGTGCAGATAGAGATAGATTTTCAATTC 449
Db 59330 ---AAGTCCAGGTTTAACTAGGTAAGTCAAGGAGGTTTATTTCAATCTTC 59386
QY 450 TTTTATCTGTCGAGACTTGCCTTTGTTTGAATAATGTAATTCATTTTCGAGA--GTTTCA 507
Db 59387 TTGTATCTGTTTATGTTGCTTGTGCTGTGATATATGATAAAATTTGAGAGGTTCTG 59446

```

TITLE
JOURNAL
COMMENT

Db	214218	ATGTCCTCCACTATACAGTGTACGAGGTCATATATATGAGGTTGGCTTTAGTATTGTTTCT	214277
Qy	722	TTTATGAACCTTGGGTGACATTTGTTTGGTGCTAGACATTAAGAAATGCAATGTCCTCT	781
Db	214278	TTTATGAATTTGGGTG-CCCTGTGTTTGGTGCATAAGTACTTAAAAATTTCAATATCCTCC	214336
Qy	782	TGGTGG-----ATTGTCCTTTGATGCGCTANGTAGTATTCTTCCCAATCTCATCTGCTT	834
Db	214337	TCTTAGTGTTTTTTTTTCTCCTTTTATAAATANGTAGTGTCCTTCCCTAGCTTCTCTGATT	214396
Qy	835	AGTTTTGGGTTTAAAGTCTA-TTAGTCAGATATTAAATGACTGTATCGGCTTGGTCTTTA	893
Db	214397	AGTTTGTATTTGAAGCCTATTTTGTGTAGATATTAAAAAGGCTATTCTGGCTTGCCTCTTG	214456
Qy	894	GGGCCATTTGCTTAGAATATCTT-----TTCCATCCTTTTACICTAAGGTGATGCTCAT	947
Db	214457	GGTCCATTCACATTAATAATCATTTGTCCATCTTCCCTGCCGCCCAAGATGATGCTCAT	214516
Qy	948	CC-----ATGGTAGGTTGCTTTTTTGGATGCAGCAGTAGATGGATCTTTGTTTTCATATCC	1004
Db	214517	CTTTAATGTTAAGATGTGTTCTTAGATGCAGCAGAAAGATGAA--TAGTTTTCATATTC	214574
Qy	1005	ATTCTGTTACCAGTATCTTTTTCTTAGAGAAATTAAGATCATTTAGTCAATTTAGTGTGAG	1064
Db	214575	ATTCGGGTTAAATAGTGTCTGTTGTAGGGGCACTTGA-----AACCATTTGCATTCGAG	214626
Qy	1065	AATTATCAATCAGCAGCTGTTTGTGATCTTGTATTCTTGCACTTGTCAAGTGTGTGTGT	1124
Db	214627	AGTTATCAATGACCAATTTTGTGATTTCCAGTTATTTTG-----TTGGTGGTGATGA	214681
Qy	1125	GT	1184
Db	214682	TGGTGTAGTGTAGTGTAGT	214741
Qy	1185	CTGTGTCTTGT	1243
Db	214742	TGGT	214801
Qy	1244	TTATTTCATATTTTCTTGAATGGGTAAACATCTTTAGATTGAAGTTTTTCTCCTAG--CC	1301
Db	214802	CTCT--GTACTTTCTTGGATATAATTGACCTCTTGAGTTTGAAGATTTGTTTCTAGTGCC	214859
Qy	1302	TCTTTTAGGCTGTGCATTTGAGATAGATATCTTTACATCTGATTTTATCTAGAAATGTC	1361
Db	214860	TATTATAGGACTGGGATTTGTAGATAGATAAAACCTAAATTTGATTTATGTGAAA----	214915
Qy	1362	TTTCTTTCTCCAACTATTGTGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTGACAT	1421
Db	214916	TTTTTTTCTCATTTATTGTGGCTGAAGTCT-----GAGCTAGGTGGGAT	214963
Qy	1422	CTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGGGCCTTCTTACATTTTGAAGTTCT	1481
Db	214964	CTGCAGACTCTTAGAAATTCATAGAACATCTGTCCAGGTTCTTCTGGCTTTTACAGTCTCC	215023
Qy	1482	ATTGGAAGTACAGGTGAATCTAATACATCTGCTTATATGTTTAAATGTCCTTTTTT	1541
Db	215024	ACTGTAAAA--CAGATATTATCTATTAGGCGTCTCTATATGCTACTTGTGCTTTTCC	215081
Qy	1542	CCCTTGCACTTTTAAATTTCTTCTTCTTCTTCTTATACCTTTTASTGATTTGATTTATGCG	1601
Db	215082	CCCTTATCTTACAATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	215140
Qy	1602	ACTGTGGGAGTTTCTTTTCCGGTCCATCTATTTTGGTGTGTTT---GTATGCTTCTTGTGA	1658
Db	215141	ACTGAGGCGAGTTCTTTTCTGGTCTGTGTTTATTTGATGTTCTGAAGTAAATCTTCTTGTGA	215200
Qy	1659	CCCTCATAGGCATCTCTTCTCAAGGTTAGGAAATTTTCTTTTGGTTTTTCTTCTGAAAA	1718
Db	215201	CTTTGATAGACATCTCTTCTTTAGATTAGAAATTTTCTTCTGTGATCTTATGAAAA	215260
Qy	1719	TATTTTCCCTGCTTTTGACCT--GCCTTCTTCCCTTCTCTATCTCT-----T	1765
Db	215261	TGTTATCTGTGCTTTGATCTGGGTTCTTCTCTCTCTATCTCTCTCTATCTCTATAGTCTTAGAT	215320

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Rhode, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Sivartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.
Direct Submission
Unpublished
2 (bases 1 to 237087)

Query Match	29.9%;	Score 680.301;	DB 1;	Length 237087;
Best Local Similarity	69.2%;	Pred. No. 0.22;		
Matches 1258;	Conservative	0;	Mismatches 477;	Indels 83; Gaps 25;
QY	306	TCATCTGGCCATTTTTAACTCAGTAGTCAGTTTGGTTTCCATAAGATTTGTGAAGT	365	
DB	75273	TTTGTCTTGAGCCATTTTCTATTCTAGTAGAGTTCCTGTCTCCATGCACATTACAGC	75214	
QY	366	TTTCTGTGTTTCTCTGTTGTGTGTATCTAGATTTAAGCTGTGGTGCAGATAGG	425	
DB	75213	TTTCTGGTGTTCCTGATA--TTGTGTGATCCAGCTTTAATCTGTGGTGCAGATAGG	75157	
QY	426	ACATAGAGTATTTATTTCAATTTGCTTTTATCTGTGAGACHTGCTTTGTTTGGAAATATG	485	
DB	75156	ACATAAGGTATTGTTTCGGTTTCTCTGTGCACTGTGTGAGACHTTCTTTGTGTCCAGTATG	75097	
QY	486	TATTCAAATTTTGG--AGAGTTTCTATAGGGTCTGCACAAGAGGTACAGCTCTTTGTGTTT	543	
DB	75096	TGGTTAATCTCGGAAAAGTTCCATGCGTTTCTGAGAAAAGGTATATTCTTTTGTGTTG	75037	
QY	544	GGTGAATAGTCTGTPAAATATCTCT--AGGTCACCTTGGTTTATGACATCATGTAGTCCA	602	
DB	75036	ATTGAATAGTTTGTGPAATACATGTTAGGTCGTGTGGTTTATGACATTAGATATCTCA	74977	
QY	603	GCAATTTCTCTGTTTGGTTTTTGTGTGAGATGACCTTAACCTGTTTGAGAGAAATGGGGTATG	662	
DB	74976	GCAATTTCTCTGTTAG--TTTTTG--GTAGGTGACCTGCTATTGTGTAGAGCAGGGTATTA	74918	
QY	663	AAGTAGCCCACTATCTGTGTGTGA--GGTCAATATGTGATTTTACGCTGTAGCTGTGTTGT	721	
DB	74917	ATGTCCTCCCACTATCAGTGTACGAGGGTCATATATAGGGTTGGCTTTAGTATTGTTCT	74858	
QY	722	TTTATGAACTTGGGGAACATTTGTTGGTGCATAGACATTAAGAATTCGATCTCTCT	781	
DB	74857	TTTATGAATAATGGGTG--CCCTGTGTTTGGTGCATAAGTACTAAAAATTTCAATATCCTCC	74799	
QY	782	TGGTCG-----ATTCTTCCTTTGATGCGCTATGTAGTATTCTTCCCAATCTCATCTCGTT	834	
DB	74798	TCCTAGTGTGTTTTTCTCCCTTTATAAATGATGATGTCCTTCCCAGCTCTCTCGATT	74739	
QY	835	AGTTTGGGTTAAGTCTA--TAGTCAGATATTAATTAAGCTGTATCGCGGTGCTCTTA	893	
DB	74738	AGTTTGTATTGAAGCCTATTTTGTAGATATTAATTAAGCTGTATCTGCGCTTGGCTCTTG	74679	
QY	894	GGGCCATTTGCTTAGAATATCTT-----TTCCCATCCTTTTACTCTAAGGTGATGCTCAT	947	
DB	74678	GGTCATTCACCTATAAATPATCATTTGTCCATCCTTTCCCTGCCCCCAAGATGATGCTAT	74619	
QY	948	CC--ATGCTAGGTTGCTTTTTTGGATGCAAGTAGTAGGATGGATCTTGTTTTCATATCC	1004	
DB	74618	CCTTAAGTTAAGATGTCTTCTTAGATGCAAGAGAAGATGAA--TAGTTTTCATATTC	74561	
QY	1005	ATTCTGTTACCCAGTATCTTTTTCTAGAGAAATTAAGATCATTTAGATCTATGATGTTGAG	1064	
DB	74560	ATTCCGTTAAATAGTGTCTGTTGTAGGGCAGATTGA-----AACCAATGACATTGAG	74509	

Qy	1065	AATTATCAATGAGGAGTGTGTGTGGAAATCTGTGTAATCTTGACCTCTGTGGAAGTGTGTGTGT	1124
Db	74508	AGTTATCAATGACACAGTATTTGTGTGATTTCCAGTTATTTTG-----TTCTGGTGGTGATGA	74454
Qy	1125	GT	1184
Db	74453	TGGTGGTAGTGTCTAGTGTCTAGT	74394
Qy	1185	CTGTGTCTGTGTGTGTGTGTCTCTCCCTCTTTTGATTTT-TGGCCTGGAAATTAATTTA	1243
Db	74393	TGGT	74334
Qy	1244	TTATTTCATATTTCTTTGAAATGGGTAAACATCTTTAGATTGAAGTTTTTCTCCCTAG--CC	1301
Db	74333	CTCT--GTACTTCTTTTGGATTAATTTGACCTCTTGAGTTTGAAGATTTGTTTCTAGTGCC	74276
Qy	1302	TTCTTTAGGTCTGCATTTGAAGATAGATATCTTTTACATCTCAATTTTATCTTAGAANGTC	1361
Db	74275	TATTTATAGGACTGGAATTTGTGTAGATAGATAAAACCTAAATTTGATTTTATTTGTGAAA----	74220
Qy	1362	TTTCTTTCTCCAACTATTTGTGACAGAAAGTTTCTTAAAGTCAGTAGTCTGGCCTGACAT	1421
Db	74219	TTTTTTTTCTCCATTTATTTGGGTGAAAGTCT-----GAGTCTAGGTGGGAT	74172
Qy	1422	CTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGGCGCTTCTTACATTTTGTAGTTTCT	1481
Db	74171	CTGCAGACTCTTAGAATTCATAGAAACATCTGTCCAGTCTCTCTGGCTTTTACAGTCTCC	74112
Qy	1482	ATTGGGAAAGTCAGGTGTAATCTAAATACATCTGCCTCTTATATGTAAATTTGGTCTTTTTT	1541
Db	74111	ACTGTAAAA--CAGATATTAATCTATTAGGCGTGTCTCTATATGCTACTTGTGCTTTTTCC	74054
Qy	1542	CCCTTGACATCTTTTAATTTCTTTCTTTGTCTCTATACTTTTGTAGTGAATTTGATTTATATGC	1601
Db	74053	CCCTTATATCTTACAAATTTCTTTCTTTGTT--TGTATGTTTATTTGTTTGAATTTATATGT	73995
Qy	1602	ACTGTGGGAGTTCCTTTTCCGGTCCAAATCTATTTGGTGTTTT---GTATGCTTCTCTGA	1658
Db	73994	ACTGAGGCGAGTTCTTTCTTCTGTGTCTGTCTTATTTTGTATGTTCTGAAGTAAACCTCTTGA	73935
Qy	1659	CCTTGATAGGCACTCTTCTCAAGTTAGGAAATTTTCTTTTGTGTTTTTCTTTGTTTCTTGAAA	1718
Db	73934	CTTTGTAGACATCTCTCTTTTAGATTAGAAATTTTCTCTGTGATTTCTATTGAAAA	73875
Qy	1719	TATTTTCCCTGCTTTTGACCT-GCCTCTTCCCTCTCTATCTCT-----T	1765
Db	73874	TGTTATCTGTGCTTTTGATCTGGGTTCCTCTCTCTCTCTATTCCTACTAGTCTTAGAT	73815
Qy	1766	TGGTTTTTGCATAGTGTCTCTGGCTTCTCTGGATGTTTTTATGCCGTGAATTTATTTAGACTT	1825
Db	73814	TTGGTCTTTCCACCTGTATTGAGATTTCTGTAACACTTTGTAGCTGAAGCTTTTATAGATTT	73755
Qy	1826	AACATTTCTTTTGACCAGGTATCCATTTCTTCTATCTTGTCTTCACTGCCCTGAGATTCT	1885
Db	73754	AACATCTCATTTGACTGAATAGTCCATTTCTTCAACATGTCTTCAAGCTTGTGATTTCT	73695
Qy	1886	CTCTTCTATCTCTGTATTTCTGTGAGTGAGGCTGTCTGTAGGTTTCCGTGTTGGGTTCTT	1945
Db	73694	TTCTTCTGTGCTTTGTATTTCTTTTGGTGAACCTTGGCTATCAGGCTCCCCTTCAGGTTCT	73635
Qy	1946	AATTTTTTTCATTTCCAGATTTCCCTCAGTTTGGGTTTTGTTTATTAATTTCTATTTCCACT	2005
Db	73634	AAATTTTTTCAPTTCCAGATTACCCTCACTTGTGAATTT--CCTAATTGACTCTATTTCCACT	73576
Qy	2006	TTCAGGTCTCGAAATGTTTTACTCATTTTCTCTCCAGTATTTTACATTTTTCATAGGTTTCT	2065
Db	73575	TTCAGGTCTCGAACTGTTTTATTTCA--TTCCTTTCAAAGGCCATGATTTTTCATGATTTCT	73517
Qy	2066	TTAATGGAATTTAATCAAT	2083
Db	73516	TTAAGAGACTTTTGTCTT	73499

RESULT 44	TITLE
AC123982	JOURNAL
LOCUS	REFERENCE
DEFINITION	AUTHORS
ACCESSION	TITLE
VERSION	JOURNAL
KEYWORDS	REFERENCE
SOURCE	
ORGANISATION	
REFERENCE	
AUTHORS	

AUTHORS Rat Genome Sequencing Consortium.
 TITLE Direct Submission
 JOURNAL Submitted (06-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 COMMENT On Nov 6, 2002 this sequence version replaced gi:23907833. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Center project name: GYOK
 Center clone name: CH230-341M21
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 195499 bases at least Q40
 Consensus quality: 197131 bases at least Q30
 Consensus quality: 198079 bases at least Q20
 Estimated insert size: 203971; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 261378: contig of 261378 bp in length.
 Location/Qualifiers
 1..261378
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-341M21"

misc_feature
 1..1697
 /note="wgs_end_extension
 clone_end:77"
 53300..55403
 /note="wgs_end_extension
 clone_end:77"
 complement(58350..59206)
 /note="clone_boundary
 clone_end:77
 site:
 end_sequence:BZ188858"
 complement(258171..258994)
 /note="clone_boundary
 clone_end:Sp6
 site:
 end_sequence:BZ188859"

misc_feature
 Query Match 29.9%; Score 680.301; DB 1; Length 261378;
 Best Local Similarity 69.2%; Pred No. 0.23;
 Matches 1258; Conservative 0; Mismatches 477; Indels 83; Gaps 25;

306 TCATCTTGGCTCATTTTAACTCAGTAGTGAGTGTGTTGGTTTCCATAAGTTTGAAGT 365
 95927 TTTGTCCTGAGCCATTTTTCATTCACATGAGTGTCTCTGCTCTCCATGCATCAAC 95986
 366 TTTCTGTGTTTCTGTGTTTGTGTTTATCTAGATTTAAAGCTGTGGTGGTCAGATAG 425
 95987 TTTCTGTGTTTCTGTGATA---TTGTTGGTATCCAGCTTTAACTGTGGTGGTCAGATAG 96043
 426 ACATAGATATTTTCAATTTCTTTTATCTGTGCAGACTTCTTTGTTTGAATATG 485
 96044 ACATAGATATTTTCTGTTTCTTCTGATCTGTTGAGACTTTCTTTGTGTCAGATAG 96103
 486 TATTCATTTTGG--AGAGTTTTCATAGGGTGTGTCAGAGAAGTACAGCTTTTGTGTTT 543
 96104 TGGTTAATCTCGGAAAAGTTCATCGGTTCTGAGAAAAGGTATATTTCTTTGTGTTG 96163
 544 GGTGAATAGTCTGTAAAATCTCT--AGGTCCACTTGGTTTATGATGATCAGTTAGTCCA 602
 96164 ATTGAATATGTTTGTAAATACATGTTAGTCTGTGTGGTTTATGACATTAGATATCTCA 96223
 603 GCATTTCTGTTTCTGTTTCTGTTTCTGATGACTTAACCTGTTGGAGAGATGGGTATTG 662
 96224 GCATTTCTGTTTAG--TTTGTGTAGTGGTCTGCTTATTTGGTGGTGGTATTATTA 96282
 663 AAGTAGCCCACTATCTGTGTGTGA--GGTCAATGATGATTTTAGCTGTAGCTGTGCTTGT 721
 96283 ATGTCTCCCACTATCAGTGTACGAGGCTCATATATGAGGTTGGCTTTAGTATTGTTCT 96342
 722 TTTATGAATCTGGTGCACATTTGTTTGTGTCATGACATTAAGATTCGAATGCTCTCT 781
 96343 TTTATGAATCTGGTGC--CCCTGTGTTTGTGTCATAGTACTAAAATTTCAATATCTCTC 96401
 782 TGTGTTG-----ATTTCCTTTGATGCCCTATGTAGTATTCTTCCCAATCTCATCTGCTT 834
 96402 TCTTAGTGTGTTTCTCTCTTTTAAATATGATGAGTGTCTTCCCTAGCTCTCTGATT 96461
 835 AGTTTGGGTTTAAAGCTA--TTAGTCAGATATAAATGACTGTATCGGCTTCTCTTA 893
 96462 AGTTTGTGTTGAGCTATTTTGTAGATATAAATAGGCTATTCTGGCTTCTCTTGT 96521
 894 GGGCCATTTGCTTAGAATATCTT-----TTCCATCTTTTACTCTAAGGTGATGCTAT 947
 96522 GGTCCATTCATTAATATCATTTGTCCTTCCCTGCCCCCAAGATGATGCTAT 96581
 948 CC---ATGTAGTGTCTTTTGTGATGAGCAGTAGGATGATCTTCTGTTTATATCC 1004
 96582 CCTTAATGTAGATGTTTCTTAGATGACAGAAAGATGAA--TAGTTTTCATATTC 96639
 1005 ATTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCAATGATGTTAG 1064
 96640 ATTCTGTTAATAGTCTGTTGTAGGGCAGTTGA-----AACCATGACATGAG 96691
 1065 AATTATCAATGACAGTGTGTTGGAATCTTGTATCTTTGCACTGTGTGAAGTGTGTGT 1124
 96692 AGTTATCAATGACAGTATTTGTTGATTTCCAGTATTATTG-----TTGTGGTGGTATGA 96746
 1125 GT 1184
 96747 TGT 96806
 1185 CTGTGTCTTGT 1243
 96807 TGT 96866
 1244 TTTATCAATTTTCTTGAATGCGGTAAACATCTTTAGATTTAGTTTCTCTCTAG--CC 1301
 96867 CTCT--GTACTTTCTTGGATATAATGACCTCTTGAGTTGAAGATTTGTTTCTAGTGCC 96924
 1302 TTTCTTAGTGTGTGATTTGAAGATATTTCTTACATCTCATTTTATCTTAGAATGTC 1361
 96925 TATTATGACATGATTTCTAGATAGATAAAACCTAAATTTGATTTTATTTGTGAAA--- 96980
 1362 TTTCTTTCTCAACTATTGTGACAGAAAGTTTTCTTAAGTGAGTAGTGTGCGCTGACAT 1421

Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GADK
 Center clone name: CH230-3B9
 ----- Summary Statistics
 Assembly program: Atlas;
 Consensus quality: 251930 bases at least Q40
 Consensus quality: 253189 bases at least Q30
 Consensus quality: 254147 bases at least Q20
 Estimated insert size: 263100; sum-of-contigs estimation
 Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 272926: contig of 272926 bp in length
 * 272927 273026: gap of unknown length
 * 273027 274459: contig of 1433 bp in length.
 Location/Qualifiers
 1. .274459
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-3B9"
 1. .2009
 misc_feature
 clone="wgs_end_extension
 clone end:17"
 complement(4162..4370)
 /note="clone_boundary
 clone end:17"
 site:EcoRI
 end_sequence:BH307336"
 Query Match 29.6%; Score 674.102; DB 1; Length 274459;
 Best Local Similarity 63.4%; Pred. No. 0.26;
 Matches 1262; Conservative 0; Mismatches 644; Indels 83; Gaps 19;
 QY 270 TTCCCTCTCTCTCAACACTTCTATTTCTTGATTTCTATCTTGGCTCATTTTAACTC 329
 DB 133594 TTCAATTGAATCTAGAAAGCCCTTAATGCTTTATTTCTTTTGACCCACTTGTCTATTG 133653
 QY 330 AGTAGTGAGTTGTTGGTTTCCATAAGTTTGTAAGTTTCTTGTTGTTTCTGTTCTGTTGTTG 389
 DB 133654 AGTAGAAAGTTGTTGAGCTTCCATGAGCTGTAGGCTTTCTGTTCTGTTCTGTTGAG 133713
 QY 390 TTGTTATCTAGATTTAAGCTGTGGTGCAGATGAGCATAGAGTATTTTCAATTGTC 449
 DB 133714 T-----CCAGCTTTTAAATCCATGTTGGTCTGATAGGTACAAAGGGGGTATTTCATTTC 133768
 QY 450 TTTTATCTGTCAGACTTGTCTTTGTTTGAATATGATTCATCAATTTTG-GAGAGTTTCAT 508
 DB 133769 TTGATCTGTTGAGACTTGTCTTTGCTAGCTAGTATATGTTAAATTTGGAGACAGTTTCAT 133828
 QY 509 AGGGTGTGACAAGAAGTACAGTCTTTGTTGTTTGGTGAATAGTCTGTAAATATCT-C 567
 DB 133829 GAGGTGCTGAGAAGAAGGTATATCTTTTGTCTTTGAGTGAATGTTTGCAGATATCTGT 133888
 QY 568 TAGGTCCACTTGGTTTATGATCATAGTTAGCTCCAGCATTTCTGTTTCTGTTTGT 627
 DB 133889 TAGGTCCACTTGGTTTATGATCATAGTTAGCTCCAGCATTTCTGTTTCTGTTTGT 133947
 QY 628 GAGATGACCTAATCTGTTGGAGAGATGGGGTATTGAAGTAGCCCACTATCTGTGTG- GA 686
 DB 133948 TAGATGACCTGACCAATGGTCATAGTGAATGTTGACGTCTCCCACTATTAAATGATGGT 134007

QY 687 GGTCAATATGTGATTTTAGCTGTAGCTGTGCTGTTTATGAACCTTGGGTGACATTTGT 746
 DB 134008 GTTCAATGTGTGATGTAACCTTTTAGTAAAGCTGCTTTCACACATGTGACTGCTTTGAGT 134067
 QY 747 TTGGTGATAGACATTAAGAAATTGCAATGTCCTCTGGTGGATTTTCCTTTGANGCCTAT 806
 DB 134068 TTTGGGTATGAATGTTTCAAAAATGAGACATTTATCTTGAATGATTTTACCTTTGATGAGTG 134127
 QY 807 GTAGTATTTCTCCCAATCTCATCTGCTTACCTTTGGGTTTAAAGTCTATTAGTCAGATATT 866
 DB 134128 CAAAGTGTCTTCCGGTTTATTAATTTTGGTGGAAATTTATATTATTCATAT 134187
 QY 867 AAAATGACTGTATCGGCTTCTTCTTAGGCCCAATTTGCTT-AGAATATCTTTTCCATCT 925
 DB 134188 TAGAATACTACTTCTTCTTCTTAGGTTTATTTGCTTGAATAAGCTTTTCCAGCCC 134247
 QY 926 TTTACTCTAGGTGATGCTCTATC---CATGGTAGGTGCTCTTTTGGATGAGCAGTAG 982
 DB 134248 TTTACCTCTGAGGTAATGCTATCTTTGATGTTGAAGTGTGTTTCTTGTATGACGCAAAAT 134307
 QY 983 GATGGATCTTTGTTTTCATATCCATCTGTTTACCCAGTATCTTTTCTAGAGAAATTAAGA 1042
 DB 134308 GA-----ATGTTCAAAATCCATTTCTCATAGGCTGTGCTTTTATTAGGGAGTTGAG- 134358
 QY 1043 TCATTGAGTCAATGATGTTGAGAAATTAACAAGAGCAGTGTGTTGGATCTTGTATTCT 1102
 DB 134359 -----TTCAATTGATTTAGAGATATTAATAAATCAATGATTTGTTAAATTTCTGTTATT 134411
 QY 1103 TGCACCTGTGAAGT 1162
 DB 134412 TGTGGT 134471
 QY 1163 GT 1222
 DB 134472 NNN 134531
 QY 1223 TTTTTCGSC---CTGGAAATTTATTTATTCATATTTTCTTGAATGTTGGTGAACATCTTT 1278
 DB 134532 GGTTTTGTGATGTGAATTTATTTCTCTGTGTGTTTTCATAGATGTAGCTAATCCCATTT 134591
 QY 1279 AGATTGAAGTTTTCCTCCTAG--CCTTCTTTAGGCTGCAATTTGAAGATAGATATTCTTT 1336
 DB 134592 GGGTTGGAGTTTTCCTCTTAGTATCTTTTGTAGGCTGATATGTAGAGATAGTGTTT 134651
 QY 1337 ACATCTGATTTTATCTTAGAATGCTTTCTTTCTCCAACTATTTGACAGAGAAGTTTTC 1396
 DB 134652 AAAATTTGTTTATCATGAGATATCTTATTTTCTCCATTTATGGTAAATGAAAGTTTTC 134711
 QY 1397 TAAGTCAGTAGTCTGGCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCA 1456
 DB 134712 TTGGTATATAGTCTGGGCTGACATTTTGTGATTTCTTAAAGACTTTAAGACATCTGTCTA 134771
 QY 1457 GGGCCTCTTACATTTTGAATTTCTATTTGGAAGTCAAGTGTAAATCTAAATACATCTGC 1516
 DB 134772 GGGCCTGCTGTCCCTTAGAGTCTCT-TTTGAGAAATTTGGGTATAATCTTAAATAGCTAC 134830
 QY 1517 CTTTATATGTAAATGGTCTTTTCCCTTGTGATCTTTTAAATTTCTTTCTTTGTTCTAT 1576
 DB 134831 TTTTATATACACTTGGACTTTCTCCACT----TACGTTTTTAATATTTCTTTTCTTCTGT 134886
 QY 1577 ACTTTTAGTGAATTTGATTTATGCACTGTGGGAGTTTCTTTTCCGGTCCAAATCTATTT 1636
 DB 134887 ATACTTAATGTTTGAATGATTTATGTCAGGAGGAGTATTTCTTTTGGGTCCAACTATTT 134946
 QY 1637 GGTGTTTCTGATGCTTCTTGTACCTTGATAGCATCTCTTTCTCAAGGTATAGGAATTT 1696
 DB 134947 GGCATTTATAGCTTCTTGTGCAATTCAGAACATCTCTTTCTTTTAAAGTTAGGGAAATTT 135006
 QY 1697 TCTTTTGGTTTCTTGAATAATTTTCCCTGCTTTTTCGACCTG----CCTTCTTCCCT 1752
 DB 135007 TCTTCTATGACTTGTGTAATAATTTTCTGGACCTTTTGAGCTAAGAAATCTTCTCTCTCT 135066
 QY 1753 TCTCTATTCCTTTGGTT-----TTTGATAGTGTCTCTGGCTTCTCTGGATGTTTTATGC 1807

```
Db 135067 TTCTATTATCTAGGCTTAATCTTTTATGGTTTCCCAAAATTCCTGAAATGTTTGTGT 135126
QY 1808 CTGGATTATTATAGACTTAACATTTTCTTTGACCAAGGTATCCATTCTTCTATCTTGTCTC 1867
Db 135127 CAGAACTTTTAAATTTACATTTTCTTTAACTGATGATCAACTTCTTCTACTGAAT 135186
QY 1868 TTCACCTGCTGAGATCTCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGA 1927
Db 135187 TTCTATGCTGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 135246
QY 1928 GGTCTCTGTTGGTCTCT-TAATTTTCTATTTCCAGATTTCCCTCAGTTTGGGTTTGT 1986
Db 135247 TGATTCCTGTTGCTTAICTAGATTTTCTAICTCCAGGATTCCTCAGTTTGTGTTTCTT 135306
QY 1987 TATTAATCTATTTCCACTTTCCAGTCTCTGAAATGTTTACTCATTTTCTCTCCAGTATT 2046
Db 135307 TACTGCTTCTATTTTCATTTTCAATTAACAGTTTATTTATTTTC--CTTCAACCAATTGAT 135364
QY 2047 TACATTTTCATAGGTTTCTTTAATGGAATTTATTCATTTCTCTCTTCAAGGACCTTTTATGA 2106
Db 135365 GACATTCCTCTGTGATTTCTTTAAGAAATCCATTCATTTCTCTCTTTAAAGAGGAT----- 135418
QY 2107 ATTCATAAAATGATGTTAAGTTCCTTGGCTTGTGCTTCCAGCTATGTTGCAATCT----- 2161
Db 135419 -----GTGAGGTCAATGCTTATGCTTATGCTTCAAGTTGCTAAATATCTAGG 135462
QY 2162 ----CAGGGCCTATGCTAATAGGTTTTHAGCAGGACATATTTGCTCTGTTGTTATTTGTC 2217
Db 135463 GTTCTTGGAGTACTGTAATGCACTCTCTGTTGGTGCATGCTGCTCTGTTCTTCTGTTGAT 135522
QY 2218 TGTGTTTTT 2226
Db 135523 TGTGTTCTT 135531
```

Search completed: August 24, 2004, 16:54:55
Job time : 11290 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 06:30:30 ; Search time 2253 Seconds
(without alignments)
3.614 Million cell updates/sec

Title: US-10-664-775-4
Perfect score: 2279
Sequence: 1 gatcaactctctagtgaag.....ttgtaattcttaggtgctgat 2279

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 0.5

Searched: 39 seqs, 1786346 residues
Total number of hits satisfying chosen parameters: 78

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : rng4.seq*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	591.499	26.0	19619	1 ABK47192	Mouse Contig 1A co
C 2	588.5	25.8	44576	1 AAJ61522	Cosmid CVO14 conta
C 3	588.3	25.8	29604	1 AAX83005	Partial mouse WRN
C 4	585.1	25.7	173810	1 ABN85752	Mouse chromosome 1
C 5	578.299	25.4	173810	1 ABN85752	Mouse chromosome 1
C 6	578.2	25.4	8048	1 ABN85760	Toxicity-related g
C 7	578.2	25.4	8048	1 ADB52444	Primary rat hepat
C 8	563.399	24.8	51259	1 AAX83007	Partial mouse WRN
C 9	555.699	24.4	47115	1 ADA02627	Mouse Flt3 carcino
C 10	555.699	24.4	47115	1 ADB72365	Mouse Flt3 gene.
C 11	525.898	23.1	96597	1 ADA02501	Mouse Bach2 carcin
C 12	525.898	23.1	96597	1 ADB72239	Mouse Bach2 gene.
C 13	521.797	22.9	51259	1 AAX83007	Partial mouse WRN
C 14	486.098	21.3	9330	1 AAD51986	Mouse Cyp3A11 gene
C 15	486.098	21.3	12275	1 AAD51995	Mouse Cyp3A11 prom
C 16	475.598	20.9	3449	1 AAF92531	Rat T2R04 nucleoti
C 17	453.098	19.9	96594	1 AAD85257	Mouse Ptpkr genom
C 18	453.098	19.9	96595	1 ADA02777	Mouse Ptpkr gene.
C 19	453.098	19.9	96595	1 ADB72515	Mouse Ptpkr gene.
C 20	381.798	16.8	41637	1 ADB02837	Mouse Map3k8 carc
C 21	381.798	16.8	41637	1 ADB72575	Mouse Map3k8 gene
C 22	381.798	16.8	41637	1 ADC85316	Human Mef2c coding
C 23	328.699	14.4	2920	1 AAX35043	Murine cDNA isolat
C 24	323.899	14.2	16442	1 AAX83006	Partial mouse WRN
C 25	302.299	13.3	41400	1 AAX77189	Mouse BAC279 clone
C 26	288.196	12.6	46765	1 AAS99306	DNA encoding Aldeh
C 27	280.299	12.3	1632	1 ADB59208	Toxicity-related g
C 28	278.099	12.2	3903	1 ADB59149	Toxicity-related g
C 29	277.897	12.2	70665	1 ABT10716	Human breast cance
C 30	269.097	11.8	56583	1 AAX35003	Human adenosine re
C 31	269.097	11.8	56583	1 AAF21125	Human low adenosin
C 32	269.097	11.8	56583	1 ABZ96819	Human nucleic acid
C 33	268.097	11.8	50000	1 AAG64139	Nucleotide sequenc

ALIGNMENTS

RESULT 1
ABK47192/c
ID ABK47192 standard; DNA; 19619 BP.

AC ABK47192;

DT 05-JUN-2002 (first entry)

XX Mouse Contig 1A containing a fragment of the Ots1-B7 gene.

ds; nucleic acid library; immune response; asthma; DC-SIGN;
airway hyperresponsiveness; bronchoalveolar manifestation;
signature sequence; SS; chronic obstructive pulmonary disease; COPD;
allergic disease; rhinitis; atopic dermatitis; urticaria;
autoimmune disease; multiple sclerosis; inflammatory bowel disease;
allograft rejection; infectious disease; Ots1-B7.

OS Mus sp.

XX WO200214366-A2.

FN 21-FEB-2002.

PD 16-AUG-2001; 2001WO-NL000610.

PF 16-AUG-2000; 2000EP-00202867.

PR (UYUT-) RIJKSUNIV UTRECHT.

PI Groot PC, Van Bergenhegouwen BJ, Van Oosterhout AJM;

DR WPI; 2002-241888/29.

XX P-PSDB; AAU88024.

XX Nucleic acid library comprising genes which are capable of initiation,
progression and suppression of an immune response, especially an immune
response observed with airway hyper-responsiveness of asthma.

PS Example 10; Fig 9; 120pp; English.

XX The invention relates to a nucleic acid library comprising genes or their
fragments which are capable of modulating an immune response observed
with airway hyperresponsiveness and/or bronchoalveolar manifestations of
asthma. Also included are a method for modulating an immune response of
an individual comprising modulating a gene comprising a nucleic acid at
least functionally equivalent to a nucleic acid identifiable by a
signature sequence (SS) given in the specification such as R1-SO-R1-All,
StO1-A10, SvO2-1-C11, StO1-A12, and R1-SO-R1-B7, a substance (for use as
a medicament) capable of modulating a gene comprising a nucleic acid at
least functionally equivalent to a nucleic acid identifiable by SS and
the use of a proteinaceous substance derived from a nucleic acid at least
functionally equivalent to a nucleic acid identifiable by SS for the
production of an antagonist (for use as a medicament) against the
substance. The antagonist and substance are useful for the treatment of
an immune response observed with airway hyperresponsiveness and/or
bronchoalveolar manifestations of asthma. The method is useful for

Human cDNA differe
Human reproductiv
Mouse dectin-2 gen
Mouse tyrosinase g
Mouse tyrosinase g
Mouse tyrosinase g
Human KCNQ5 (KCN6q
Chemically pretrea
Human breast cance
Human KCNQ5 (KCN6q
Mouse Ptpkr genom
Mouse Ptpkr carcin

modulating the above immune response, where the gene encodes a gene product capable of modulating the immune response. The substance is useful for treating an immune response, particularly asthma, chronic obstructive pulmonary disease (COPD), allergic diseases (rhinitis, atopic dermatitis, urticaria), autoimmune diseases (e.g. multiple sclerosis), inflammatory bowel disease, allograft rejection and infectious disease. The present sequence is a genomic DNA sequence (Contig 1A) comprising part of the gene sequence of DC-SIGN (a signature sequence also known as O6S1-B7)

XX 19-JUN-2000 (first entry)
 XX Cosmid CVO14 containing rat vasopressin and oxytocin locus.
 DE Oxytocin expressed sequence tag; 5'-OT EST; obesity; fertility; male;
 XX transgenic animal; human late onset obesity; late onset visceral obesity;
 KW male infertility; wasting; anorexia; cachexia; malabsorptive state;
 KW catabolic state; inflammatory condition; Crohn's disease; AIDS wasting;
 KW burn; cancer; bone disease; vasopressin; oxytocin; ss.
 XX Rattus sp.
 OS WO200009686-A1.
 XX 24-FEB-2000.
 PD
 XX 12-AUG-1999; 99WO-GB002658.
 XX 12-AUG-1998; 98GB-00017566.
 PR 06-MAY-1999; 99GB-00010522.
 XX (MED-) MEDICAL RES COUNCIL.
 PA Robinson ICAF, Stoye JP, Flavell D, Wells SE, Le Tissier P;
 PI WPI; 2000-224331/19.
 XX New anti-obesity polypeptide useful for treating obesity or infertility
 PT in mammals.
 XX Claim 9; Page 129-154; 162pp; English.
 XX The present sequence represents cosmid CVO14, which contains the rat
 CC vasopressin and oxytocin locus. The specification describes 5'-OT-EST
 CC (oxytocin expressed sequence tag) polypeptides. The 5'-OT EST gene is
 CC involved in the control of obesity and fertility in males. 5'-OT EST
 CC nucleic acids are useful for producing transgenic animals. The transgenic
 CC animals created serve as a model for human late onset obesity and other
 CC related disorders and are also used for identifying the genetic cause of
 CC obesity. Compounds which modulate 5'-OT EST expression or activity are
 CC useful in the treatment or modulation of late onset visceral obesity or
 CC male infertility particularly in the disorders related to these
 CC conditions such as wasting, or anorexia, or cachexia associated with
 CC prolonged illness, or malabsorptive states or catabolic states associated
 CC with other diseases such as inflammatory conditions, Crohn's disease or
 CC AIDS wasting, or burns, or cancer, or bone disease
 XX SQ Sequence 44576 BP; 12157 A; 10993 C; 10857 G; 10569 T; 0 U; 0 Other;
 Query Match 25.8%; Score 588.5; DB 1; Length 44576;
 Best Local Similarity 62.7%; Pred. No. 1.2;
 Matches 1173; Conservative 0; Mismatches 560; Indels 137; Gaps 18;
 QY 292 TCTATTCTTGATTTCTATCTGGCTCATTTTACTCAGTAGTAGTGTTGTTTCC 351
 DB 16417 TTCTTTCTTTATTTCTTCCCTGACAGGTTATCATTTAGTAGAGCATTTTCC 16358
 QY 352 ATAAGTTTGTAGTTTCTGTTTCTGTTTGTGTTGTTATCTAGATTAAAGCTGT 411
 DB 16357 AAGTATATGGGCATTC-----TTCCTTGATTGTTATTGAAGACCACTTTAGGCCGT 16304
 QY 412 GGTGGTCAGATGGACATAGAGTATTATTTCAATGTCTTTTATCTGTGAGACATTTGTT 471
 DB 16303 GGTGGTCCGATAGCATGATGGGATTATTTCTATCTTTCTGTACCTGTGGGCCCGTT 16244
 QY 472 TGTTTTGAATATGTTATCAATTTTGGAG--AGTTTCATAGGTTGCTGACAGAAGGTAC 529
 DB 16243 TTGACCAATTATGCTGCAATTTTGGAGAAGTAGTACCATTGAGAGAGGTAT 16184
 QY 530 AGCTTTTGTGTTGG-TGAAATAGTCTGTAATAT-CTCTAGGTCCACTTGGTTATGA 587
 DB 16183 ATCCCTTTTGTAGTAGAATGTTCTATAATATCCGTTAAGTCCCTTTGGCTCATGA 16124

QY 588 CATCAGTTAGTCCAGCATTTCTCTGTTTGGTTTTTGTGAGATGACCTAACHTGTGGA 647
 DB 16123 CTCTCTTAGTCTGTACATCTCTCTGTTT-AAATTCCTGTTTCCATGATCTGTCCATGAT 16065
 QY 648 GAGAAATGGGTATTGAAGTAGCCACATCTCTGTGTGTAGGT-CAATATGTGATTTTTCAGC 706
 DB 16064 GAGAGTGGGTGTGAATCTCCACTATTATTGTGTGAGGTGCAATGTGTGTTTTCAGC 16005
 QY 707 TGTAGTGTGCTGTTTGTATGAACCTGGGTGACATCTGTGTTGGTGCATGACATTAAGA 766
 DB 16004 TTTAGTAAGGTTTCTTTTACATATGTAGTGCCTTGTATTGGGGCATAGATATTAGG 15945
 QY 767 ATTGCAATGCTCTCTGCTGGA-TTTTCCCTTGTATGCTATGTAGTATCTTCCCAATCT 825
 DB 15944 ATTGAGATTCATCTTGTGTGATTTTCCCTTGTGATGATGAGTGTCTCTTCTTAATCT 15885
 QY 826 CATCTGCTTAGTTTGGTTTAAAGTCTA-TTAGTFCAGATATTAAATGACTGTATCGGCT 884
 DB 15884 TTTTGTATGACTTTTAGTAAATATGATTTTATTTGATATTAGATGGCTACTCCAGCT 15825
 QY 885 TGCTTCTAGGGCCATTTGCTTAGAATA-TCTTTTCCATCTCTTACTCTAAGGTGATGT 943
 DB 15824 TGCTTCTTCTGACATTTGCTTGGAAATTTGTTTCCAGGCTTTCACCTCTGAGTAAATGT 15765
 QY 944 CTATCCATG---GTAGTGTCTTTTGGATGAGCAGTAGGATGATCTGTGTTTCAT 1000
 DB 15764 CTGTCTTGTCTGTAGGTTGTTTCTGTAGGACAGAAATGCAGGCTCTCTGTTGCGT 15705
 QY 1001 ATCCATCTGTATCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATTTGATGT 1060
 DB 15704 ATCCAGTTTGTAAATCTATGTCTTTTATTGGGAGTTAAG-----GCCATTTATGT 15653
 QY 1061 TGAGAAATTAATCAATGACAGCTGTTTGGGATCTTGTATCTTGTCACTGTGGAAGTGTGT 1120
 DB 15652 TGAGAGATTAATGAAGATAGTATTTGCTTCCGCTTATTTATCATATTGGATGAGG 15593
 QY 1121 GT 1180
 DB 15592 TTATGTTTGTGTG----- 15580
 QY 1181 GT 1240
 DB 15579 -----CTTTCATCTGCTTTGTTTGTGCGAAGACGAT 15547
 QY 1241 TTATTTATTCATATTCTTGAATGTGGTAACATCTTTAGATTGAAGCTTTTCTCTCAGC 1300
 DB 15546 TAGTTTCTGTCTTCTCTAGGATAGCTTGCCTCTTATGTGGGCTTTACCATTTATT 15487
 QY 1301 CTTCTT--TAGGTCTGATTTGAAGATAGATATTCTTTATCATCTGATTTTATCTTAGAAT 1358
 DB 15486 ATCTTTGTAGTGTGGATTTGTAGAAAGATATTGTGTAAATTTGGTTTGTGTCATGGAAT 15427
 QY 1359 GTCTTTCTTTCCCAACTATTGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTGA 1418
 DB 15426 ATCTTGGTTTCTCATCTATGTTAATTGAGAGTTTTCCAGGATACAGTAACCTGGGCTGG 15367
 QY 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGGCCCTTCTTACATTTTGAAT 1478
 DB 15366 CATTTGTGTTCTCTTAGGGTCTGTATGACATCTGTCCAGGATCTTCTGGGCTTCATAGTT 15307
 QY 1479 TCTATTGGAAAGTGCAGGTGATTTCTTAATACATCTGCCCTTATATCTTAATTTGGTCTTT 1538
 DB 15306 TCT-GGGCAGAGGCTCTGGTGTGATTTCTGATAGGCTCTGCCCTTATATGTACTT-GACCTT 15249
 QY 1539 TTTCCCTTGTATCTTTTAAATATTCTTTCTTGTCTTATATCTTTTGTAGTATTGATTTA 1598
 DB 15248 TTTCCCTTACTGCTTTTAAATATTCTTTCTTTATTTTGTGCGTTTGGTGTGTTGACTATTA 15189
 QY 1599 TGCATGTGGGGAGTTTCTTTTCCGGTCCAAATCTATTGTTGGTGTGTTTGTATGCTTCTGTA 1658
 DB 15188 TGTGACGGGAGGTGTTTCTTTTCTGGTCCAATCTATTGGAAGTTCCTGTAGGCTTCTTGTGA 15129

Db	24550	TTCTGTTCTTGTGGCTTTTCTTCTTTTGGTTTG-----	24588
Qy	1181	GTGTCTGTGTCCTTGTGTGTCTGTCTCTCCCTCTTTTGATTTTGGCCTGGAATTAT	1240
Db	24583	-----TTGAATGAT	24591
Qy	1241	TTATTATTCAATATTTCTTGAATGTGGTAACAATCTTTAGATTGAAGTTTTCCTCCTAGC	1300
Db	24592	TACTTTCTCTGGTTGTTCTTAGGGCGTGAATTCGGTTCTCTGTATTGCTCTCTTTCTGTATT	24651
Qy	1301	CTTCTCTTT--AGGTCTGCAATTTGAAGATAGATATCTTTACATCTCATTTTATCTTTAGAAAT	1358
Db	24652	ATCCTTTGAAGGGCTGGAATTCGTGGAAAGATATTGTGGAAATTTGTTTTGTCTGTGGAAAT	24711
Qy	1359	GTCTTTCTTTCTCCAACTATTGTGCAGAAAGTTTTCCTAAGTCAGTAGTCTGCCTGA	1418
Db	24712	ACTTTGGTTTCTCCATCTATGGTAAATTGAGAGTTTGGCCTGTGTATAGTACCTGGGCTGG	24771
Qy	1419	CATCTGTAGTCTCTTTGGAGTCTGTACACAATCTGTGCAGGGCCCTTCTTACATTTTGAGTT	1478
Db	24772	CATTGTGTCTCTTTAGTTTCTGTATAAACAATCTGTCCAGGCTCTTCTGGCTTTCATATGC	24831
Qy	1479	TCTATTGGAAAAGTCAGGTGTAATTCCTAATACATCTGCCTTTATATGTAAATGGTCTTT	1538
Db	24832	TCTGGT--GAAAAGTCTGGTGTAAATTCGTATAGCCCTCTCTTTATATGTACTT--GACCTT	24889
Qy	1539	TTTCCCTTGCACTTTTAAATATCTTTCTTTGTCTATACTTTTAGTGATTTGATTAATTA	1598
Db	24890	TCTCCCTTACTGCTTTTAAATATCTATCTATTATTAGTGCAATCTGTGTCTGTATTATA	24949
Qy	1599	TGCACGTGGGGAGTTTCTTTCCGGTCCAACTCTATTTTGGTGTTTGTATGCTCTCTGTATA	1658
Db	24950	TGTGTGGGAGGAATTTCTTTCTGTGTCAAGTCTATTGGAGTTCTGTAGGCTTCTGTATA	25009
Qy	1659	CCTTGTATAGCACTCTTTCTCTCAAGGTTAGGAAATTTTCTTTTGGTTTCTTGTGAATA	1718
Db	25010	TGATCATGGCACTCTTTTTTTAUGTTTGGGAAGTTTCTTCTATATTATTGTGTGAAGA	25069
Qy	1719	TATTTTCCCTGCTTTTGAACGTG-----CCTTCTTCCCTTCCCTATATCCT	1764
Db	25070	TATTAGCTGGCCCTTTAAAGTTGAAAATCTTCAATCTCATCAATTCCTATTATCCGTAGGT	25129
Qy	1765	TTGGTTTTTGATAGTGTCTCTGGCTTCTCGGATGTTTTATGCCCTGGATATTATTAGACT	1824
Db	25130	TTGGTCTTCTCATGTGTCTCTGGATTACCTGTGATGTTTGAAGTTAGGAATCCTTTTGCAAT	25189
Qy	1825	TAACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGTCTCAGATTTC	1884
Db	25190	TTGTATTTTTCTTTGACUUGTGTGTCCATGTTCTCTATGAATCTTTCACCTGAGATTTC	25249
Qy	1885	TCTCTTCTATCTCTGTATTCTGTGTAGTGAGGCTGTCTCTGAGGTTCCGTG--TTGGGTTC	1943
Db	25250	TCTCTTCCATTTCTTTGTATTCTGTGTCTGATGCTGCATCTATGGTTCCAGATCTCTTTC	25309
Qy	1944	TTAAATTTTTTCATTTCCAGATTTCCCTCAGTTTGGGTTTGGTTTATTATTAATCTATTCCCA	2003
Db	25310	CTAGGATTTCTATCTCCAGGTTGCGCTGGTTTGGGTTTCTTTTATGTGTCTACTCTCCC	25369
Qy	2004	CTTTCAGGTCCTGAAATGTTTTACTCAATTTTCTCCCGATATT---TACATTTTTCATAGG	2060
Db	25370	CTTTTAGTCTAGTATGTTTTGTTCTATTTCATCCATCACCTGTTTGGATGTGTTTCTCTGTT	25429
Qy	2061	TTTCTTTAATGGAAATTATATTCATTTCTCTTCAAGGACCTTTTATGAAATTCATAAATGTA	2120
Db	25430	TTTCTTTAATGATTTCTACTCGTTTGGCTGTGTGTTTCTGCTTTTCTTTTAAGGGCTGTATA	25489

RESULT 4

ABN85752/c
ID ABN85752 standard: DNA: 173810 BP.



AC ABN85752;

XX

DT	14-OCT-2002	(first entry)
XX	Mouse chromosome 11	BAC clone RP23-270L8 SEQ ID NO 10.
DE	Mouse; Can 1;	antinfertility; gynaecological; infertility;
XX	premature ovarian failure; menopause;	Sartoli Cell only syndrome;
KW	BAC clone RP23-270L8;	GenBank reference AC083815; chromosome 11; ds.
XX	Mus musculus.	
OS	US2002119929-A1.	
XX	29-AUG-2002.	
XX	02-NOV-2001;	2001US-00003806.
PF	03-NOV-2000;	2000US-0245872P.
XX	(BISH//) BISHOP C'E.	
PA	(AGOU//) AGOULNIK A I.	
PA	(ZHUQ//) ZHU Q.	
FI	Bishop CE, Agoulnik AI, Zhu Q;	
XX	WPI; 2002-618953/66.	
DR	A nucleic acid molecule (I) encoding a Can 1 polypeptide used in treating	
PT	infertility.	
XX	Disclosure; Page; 45pp; English.	
PS	The invention relates to a nucleic acid molecule (I) encoding a Can 1	
XX	polypeptide. The Can 1 nucleic acid molecule is used to diagnose or treat	
CC	infertility or premature ovarian failure or Sertoli Cell only syndrome in	
CC	a mammal. The present sequence is that of a Can 1 encoding	
CC	chromosome 11 BAC clone RP23-270L8 of the invention. Note: The present	
CC	sequence is not given in the printed specification but was obtained	
CC	through the GenBank reference AC083815	
XX	Sequence 173810 BP; 50474 A; 33519 C; 33994 G; 55823 T; 0 U; 0 Other;	
XX	Query Match	25.7%; Score 585.1; DB 1; Length 173810;
XX	Best Local Similarity	62.8%; Pred. No. 0.69;
XX	Matches 1168; Conservative	0; Mismatches 569; Indels 123; Gaps 18;
QY	292	TCATTTCTTCAATTCATCTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGGTTTCC 351
DB	123634	TTTCCTTTCTTATTCCTTCCTTGACCAAGGATCATTCAGAAAGAGTGTGTTTCAGTTCC 123375
QY	352	ATAAGTTTGTAAGTTTCTGTTGTTCTGTTGTTGTTGTTGTTATCTAGATTAAAGCTGT 411
DB	123574	ATGTGAAGTTGGCTTCTCTGTTATTTATTTGTTATTGAAGAT---CAGCCTTAGTGCA 123518
QY	412	GGTGTCAGATAGGACATAGAGTATATTTCATATGTCCTTTTATCTGTCGAGACTTGCTT 471
DB	123517	GGTGCTCTGATAGGATACATAGGACAAATTCAAATATTTTGAATCTGTTGAGGCGCTGAT 123458
QY	472	TGTTTTGAAATGATGATTCAATTTTGGAGA--GTTTCATAGGCTGTCACAAGAAGGTAC 529
DB	123457	TGTGACCTATTATGTGGTCAATTTTGGAGAGGTACCATGAGTGCTGAGAGAGGATAT 123398
QY	530	AGTC-TTTGTGTTTTGGTGAATAGTCTGTAAATATCTCT-AGGTCACCTTGGTTTATGA 587
DB	123397	ATCCTTTGTTTTAGGGTAAATATGTTCTGTAGATATCTGTCAGATCCATTTGTTTCATCA 123338
QY	598	CATCAGTTAGCTCCAGCAATTTCTCTGTTTCTGTTTTTGTGTTGAGATGACCTAACTGTGGA 647
DB	123337	CTTCTGTGTAGTTTCAGTGTGTCCTGTTTGTAGTTCT-GTTTCCATGATCTGTCCATTGT-123279
QY	648	GAGATGGGGTATTGAAGTAGCCCACTATCTGTGTGTGAGGT-CAATATGTGATTTTAGC 706
DB	123278	GAAGTGTGTGTGTTGAAGTCTCCCACTATTTTGTGTGAGGTGCAATGTGCTTTGAGC 123219

Query Match	25.7%	Score 585.1	DB 1	Length 173810	
Best Local Similarity	62.8%	Prod. No. 0.69			
Matches 1168	Conservative	0	Mismatches 569	Indels 123	Gaps 18
QY	292	TCTATTTCTGATTTCTATCTTTGGCTCATTTTAACTCAGTAGTCAGTTGTTTGGTTTCC	351		
DB	123634	TTTCTTTCTTTATTCCTTCCCTTGACCAAGGTATCATTGAGAAGAGTGTGTTTCAGTTTCC	123755		
QY	352	ATAAGTTTGTAAATTTTCTGTGTTTCTGTGTTTCTGTGTTATCTAGATTTAAAGCTGT	411		
DB	123574	ATGTGAATGTGTGGCTTTCGTGTTATTTATTTTGTATTGAAAGT---CAGCCTTAGTGCA	123518		
QY	412	GGTGTCAGATAGGACATAGAGTATATTCCAAATGTCTTTTATCTGTGCGAGACTTGCT	471		
DB	123517	GGTGTCGTGATAGGATACATGGGACAAATTCAAATTTTGAATCTGTGAGGCCTGAT	123458		
QY	472	TGTTTGTAAATATGTATTCAATTTTTGGAGA--GTTTCAATAGGGTGCTGCAAGAAGGTAC	529		
DB	123457	TGTGACCTATTATGTGGTCAATTTTGGAGAAGGTACCATGAGGTGCTGAGAAGAAGGTAT	123398		
QY	530	AGTC-TTTGTGTTTGGTGAATAGTCGTGTAATATCTCT-AGGTCCACTTGGTTTATGA	587		
DB	123397	ATCCTTTTGTTTTAAAGGTPAAAATGTTCTGTAGATATCTGTACAGATCCCATTTGTTTCATCA	123338		
QY	588	CATCAGTTTAGCTCAGCAATTTCTCTGTTTCTGTTTGTGTGAGATGACCTAACTGTGGA	647		
DB	123337	CTTCTGTTAGTTTCAGTGTGTCCTGTGTTAGTTTCT-GTTTCCATGATCTGTCCATTGCT-123219	123279		
QY	648	GAGAAATGGGGTATTGAAATGACCCCACTATCTGTGTGTGAGGT-CAATATGTGATTTTAGC	706		
DB	123278	GAAAGTGTGTGTGTAAGTCTCCCACTATATTGTGTGAGGTGCAATGTGCTGTTGAGC	123219		

QY 707 TGTAGCTGCTGTTGTTTATGAACCTGGGTGACATGTTGTTGGTCATAGACATTAGA 766
Db 123218 TTTACTAAAGTTCTTTAGTGAATGGCTGCTCTGTTATTTGGAGCATAGATATTAGA 123159
QY 767 ATTGCAATGCTCTCTGTTGGATTTTCCCTTGTAGTCTAGTATTTCTCCCATCT 825
Db 123158 ATTGAGATGCTCTCTGTTGGAGATTTTACCTTTGTAGTGAATGAAGTGTCCCTCTCTCT 123099
QY 826 CATCTGCTAGTTTGGGTTTAAAGTCTAT-TAGTCAGATATTAATAAGACTGTATCGGCT 884
Db 123098 TTTTGTAGTCTTGGTTGGAGTCAATCTATCAGATATTAGATGGCTACTCTGCT 123039
QY 885 TGTCTCTTAGGCGCATTTGCTTAGAATTA-TCTTTTCAATCTTTTACTCTAAGTGTATGT 943
Db 123038 TGTCTCTCATACCATTTGCTTGGAAATTTGTTTCCAGCCTTTCACTCTGAGGTAGTGT 122979
QY 944 CTATCCAT--GGTAGGTTGCTTTTGGAGTGCAGCAGTAGTAGGATGATCTTTTTCAT 1000
Db 122978 CTATCTTTTCTCTGAGATAGTTTCTGTAGCAGCAAAATGTTGGGTCTTGTGTGT 122919
QY 1001 ATCCATCTGTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTAGAGTCATGATGT 1060
Db 122918 AGCCAGTTTGTAGTCTATGCTTTTATTTGGCGAGTTGAGA-----CCATTGATGT 122867
QY 1061 TCGAATTAATCAATGAGCAGTGTGTTGGATTCTGTTATCTTGCACTGTGTAAGTGTGT 1120
Db 122866 TAAGAGATTAAGGAAAGTAATGTTGCTTCTGTTATTTTAGTTGTTAAAGTGGCA 122807
QY 1121 GT 1180
Db 122806 TTTCTGTTCTGTGGCTGCTCTTTTAGTTTG----- 122774
QY 1181 GT 1240
Db 122773 -----TTGAGGGAT 122765
QY 1241 TTATTAATCATATTTTCTGTAATGGGTAAACATCTTTAGATTGAAGTTTTCCTTAGC 1300
Db 122764 TACCTTCTGTTTCTTAGGGCTGTGTCGGTCTGTGTTATTTGTTTCTGTTATT 122705
QY 1301 CTTCTTT--AGTCTGCAATTTGAAGATAGATATTTTACATCTGATTTTATCTTAGAAT 1358
Db 122704 ATCTTTGAAGGGCTGGATCTGTGAGAGATAAATCGTGAATTTGGTTTGTCTGGAAT 122645
QY 1359 GTCTTTCTTCTCCAACTATTGTGACAGAAATTTTCTAAGTCAGTAGTCTGCGCTGA 1418
Db 122644 ACTTTGGTTTCTCTCTATGATTAATTGAGAGTTTGGCTGGGTATAGTACCTGGGTGC 122585
QY 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCGAGGCTCTTCTTACATTTTGAGTT 1478
Db 122584 AGTTTGTGTTCTTTAGTGTCTGTATAACATCTGTCCAGGCTCTTCTGCTTTCATAGTC 122525
QY 1479 TCTATTGAAAAGTCAGGTGAATTTCTAATACATCTGCTTTATATGTTAATTTGCTTT 1538
Db 122524 TCTGTT-GAAAAATCTGTTGAATTTCTGATAGGCTTGGCTTTATATGTTACTT-GACCTT 122467
QY 1539 TTTCCCTTGCACTTTTAAATTTCTTTTCTTATCTTATCTTATCTTATGATTTGATTATTA 1598
Db 122466 TTTCCCTTACTGTTTATGATTTCTATCTTTATTTAGTGCATTTGATGTTCTGATTATTA 122407
QY 1599 TGCATGTGGGAGTTTCTTTTCCGGTCCAACTATTTTGGTGGTTTGTATGCTCTCTGTA 1658
Db 122406 TGTGTGGGAGGAATTTCTTTTCTGGTCCAGTCTATTTTGGAGTTCTGTAGGCTTTCTGTA 122347
QY 1659 CTTTCATAGGATCTCTTTCTCAAGGTTAGAGAAATTTTCTTTTGTGTTTCTTGTGAAA 1718
Db 122346 TCTTCATATGCACTCATCTTTTAGATTGGAGAGTTTCTCAATAATTTTGTGTGAGA 122287
QY 1719 TATTTTCCCTGTTTGTACCTG-----CCTTCTTCCCTTCTCTATTCCT 1764
Db 122286 TGTTTGCTGGACCTTTGAGTTGAAATCTTCAATCTCATCCACTCTTATTTATCGGTACGT 122227
QY 1765 TTGGTTTTTGCATAGTGTCTCTGGCTTCTCGATGTTTATGACCTGGATTTATTAGACT 1824

Db 122226 TTGGTCTTCTTATTGTTGTCCTGATTTCTCTGGATATTTTGATTTAGGATCTTTTTCATT 122167
QY 1825 TAACATTTTCTTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCTTCTTCTGAGATTC 1884
Db 122166 TTCCATTTTCTTTGATTTGTTGCGGATGTTCTCTATGGAATCTTCTGACCTGAGATTC 122107
QY 1885 TCTTCTATCTCTTGTATCTTGTCTGAGTGGCTTGTCTGAGGTTCTG-TTGGGTTCT 1943
Db 122106 TCTCTTCCATCTCTTGTATCTTGTCTGCTGATCTCAAACTCTATGTTCCAGATTTCTTTC 122047
QY 1944 TTAATTTTCTTCAATTTCCAGATTTCTTCACTTGTGTTTGTGTTTATTAATCTATTTCCA 2003
Db 122046 CTAGGTTTCTATCTCTAGTGTGCTGCTGTTGAGTTTCTTATTTGTGTTACTTCCC 121987
QY 2004 CTTTCAGGCTCTGAAATGTTTACTCATTTTCTCTCC---AGTATTTACATTTTCATAGG 2060
Db 121986 TTTTATGTTCTAGTAGTGTGTTTGTTCATTTCCATCCATCCCTGTTGTATGTTTTCCTCTT 121927
QY 2061 TTTCTTTATGATTTATTCATTTCTCTCTTCAAGGACCTTTTATGAATTCATAAATGTA 2120
Db 121926 TTTCTGTAAGGACTTCTACCTGTTT-GATTGTGTTTCTCTGTTTCTTTAAGGACTTGTGA 121867

RESULT 5
ABN85752
ID ABN85752 standard; DNA; 173810 BP.
XX
AC ABN85752;
XX
DT 14-OCT-2002 (first entry)
XX
DE Mouse chromosome 11 BAC clone RP23-270L8 SEQ ID NO 10.
XX
KW Mouse; Can 1; antiinfertility; gynaecological; infertility;
KW premature ovarian failure; menopause; Sertoli Cell only syndrome;
KW BAC clone RP23-270L8; GenBank reference AC083815; chromosome 11; ds.
XX
OS Mus musculus.
XX
PN US2002119929-A1.
XX
PD 29-AUG-2002.
XX
PF 02-NOV-2001; 2001US-00003806.
XX
PR 03-NOV-2000; 2000US-0245872P.
XX
PA (BISH/) BISHOP C B.
PA (AGOU/) AGOULNIK A I.
PA (ZHUQ/) ZHU Q.
XX
PI Bishop CE, AgoulNIK AI, Zhu Q;
XX
XX WPI; 2002-618953/66.
XX
XX A nucleic acid molecule (I) encoding a Can 1 polypeptide used in treating infertility.
XX
XX Disclosure; Page; 45pp; English.
XX
XX The invention relates to a nucleic acid molecule (I) encoding a Can 1 polypeptide. The Can 1 nucleic acid molecule is used to diagnose or treat infertility or premature ovarian failure or Sertoli Cell only syndrome in a mammal. The present sequence is that of a mouse Can 1 encoding chromosome 11 BAC clone RP23-270L8 of the invention. Note: the present sequence is not given in the printed specification but was obtained through the GenBank reference AC083815
XX
SQ Sequence 173810 BP; 50474 A; 33519 C; 33994 G; 55823 T; 0 U; 0 Other;
Query March 25.4%; Score 578.299; DB 1; Length 173810;
Best Local Similarity 63.2%; Pred. No. 0.74;

PF 31-JAN-2003; 2003WO-US003194.
 XX 31-JAN-2002; 2002US-00060087.
 PR 15-MAR-2002; 2002US-0364045P.
 PR 15-MAR-2002; 2002US-0364055P.
 PR 30-DEC-2002; 2002US-0436643P.
 XX (GENE-) GENE LOGIC INC.
 PA Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
 PI WPI; 2003-689530/65.
 DR Predicting a toxic effect of a compound, useful in identifying toxicity
 XX markers in liver tissues or cells for drug screening and toxicity assays,
 PT comprises preparing gene expression profile of tissue or cells exposed to
 PT the compound.
 PS Claim 1; SEQ ID NO 2986; 1156pp; English.
 XX
 CC The present invention relates to a method for predicting a toxic effect
 CC of a compound. The method comprises preparing a gene expression profile
 CC of a tissue or cell sample exposed to the compound, and comparing the
 CC gene expression profile to a database comprising SEQ ID 1-4925, where
 CC differential expression of the gene indicates at least one toxic effect.
 CC The method is useful for predicting at least one toxic effect of a
 CC compound, predicting hepatotoxicity or the progression of a toxic effect
 CC of a compound, identifying an agent that modulates the onset or
 CC progression of a toxic response, predicting the cellular pathways that a
 CC compound modulates in a cell, and identifying an agent that modulates at
 CC least one activity of a protein. The method and compositions of the
 CC present invention using a database of genes having liver toxin-induced
 CC differential expression, are useful in identifying toxicity markers in
 CC liver tissues or cells for drug screening and toxicity assays. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 8048 BP; 3068 A; 1764 C; 1532 G; 1682 T; 0 U; 2 Other;
 Query Match 25.4%; Score 578.2; DB 1; Length 8048;
 Best Local Similarity 62.04; Pred. No. 2.8;
 Matches 1160; Conservative 0; Mismatches 574; Indels 136; Gaps 17;
 QY 292 TCTATTCTTCAATTTCTATCTTGCTCTCAATTTTAACTCAGTAGTGAGTGTGTTGGTTCC 351
 DB 3329 TTTCTTTCTTATTTCTTCTTGACAGGTTATCATTTGAGTAGGCAATGTTCAATTTCC 3270
 QY 352 ATAAGTTTGAAGTTTCTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
 DB 3269 AGGTATATGTGGGATTC-----TTCCTTATTTGTTATTTGAAGACCAAGTTTGGCCCGT 3216
 QY 412 GGTGTCAGATAGGACATAGATATTTCAATTTGCTTTTATCTGTCAGACTTGGCTT 471
 DB 3215 GGTGTCAGATAGGACATAGGATTTTCTATCTTCTGACCTGTTGAGCCCGTTT 3156
 QY 472 TGTGTTGAATATATTTCAATTTTGGAG--AGTTTCATAGGTCGTGACAAAGAGTAC 529
 DB 3155 TTTGACCAATTTATATGTTCAATTTTGGAGAAAGTGCCATGAGGAGCTGAGAAGAGTAT 3096
 QY 530 AGTCTTTGTTGTTGGTGAATAGTCTGTAA--TATCTTAGGTCCACTGTTGTTATGA 587
 DB 3095 ATCCTTTTGTAGGATAGAGTTCTAAATATATCCGTTAAGTCCATTTGGCTCATGA 3036
 QY 588 CATCAGTAGTCCAGCAATTTCTGTTTCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 647
 DB 3035 CTTCCTCTATGCTGTCACATCACTGTTT-AAATTCCTGTTTCCATGATGTCCTCATGAT 2977
 QY 648 GAGATCGGATATGAGTAGCCCACTATCTGTGTGAGGT-CAATATGTGATTTAGC 706
 DB 2976 GAGATGCTGTGTTGAAATCTCCACATTTATTTGTGTGAGGTGCAATGTGTTTGGAGC 2917
 QY 707 TGTAGCTGTGCTGTTTATGAACTTGGGTGACATTTGTTGTTGGTCATAGACATTAAGA 766

Db 2916 TTTAGTAAGTTTCTTTTACGTATGTAGGTGCCCTTGTATTTGGGCATAGATATTTAGG 2857
 QY 767 ATTGCAATGCTCTCTTGTGGTGA-TTTTCTTTGATGCCCTATGTAGTATTTCTTCCAACTCT 825
 Db 2856 AITGAGAGTTCACTTTGGTGGATTTTCTTTGATGAATGAAGTGTCTTCTTCTTACT 2797
 QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTATGTCAGATATTAATAAGTGTATCGGCT 884
 Db 2796 TTTTGTATCACTTTTATGTTGGAATTTGATTTTATTTGATATTAAGATGGCTACTCCAGT 2737
 QY 885 TGCTTCTTAGGCCCATTGCTTAG-AAATATCTTTTCCATCTTTTACTCTTAAGTGTATGT 943
 Db 2736 TGCTTCTTAGCAATTTGCTTGGAAATTTGTTTCCAGCCCTTCACTCTGAGGTAGTGT 2677
 QY 944 CTATCCATG--GTAGGTTGTCTTTTGGATGTCAGCAGTAGGATGGAATCTTGTTTTCAT 1000
 Db 2676 CTGTCTTTGTCTCTGAGGTGTGTTTCTGTAGGCAGCAAGATGCAGGTCCTCGTTGTGT 2617
 QY 1001 ATCCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGT 1060
 Db 2616 ATCCAGTTTGTAACTATGACTTTTATTTGGGAGTTGAG-----GCCATTTGATGT 2565
 QY 1061 TGAGAAATATCAATCAGCAGTGTGTTGTGATTTCTTGTATCTTTGACCTTTGCAAGTGTGT 1120
 Db 2564 TGAGAGATATTAAGGAATAGTATTTATGCTTCCGTTTATATTCATATTTGATGTGAGG 2505
 QY 1121 GT 1180
 Db 2504 TTATGTTTGTGTG----- 2492
 QY 1181 GT 1240
 Db 2491 -----CTTTCATCTCTTTGTTTGTGCAAGACGAT 2459
 QY 1241 TTATTTATCAATTTTCTTGAATGTTGGGTAACTTTTGTAGATTTGAAGTTTCTTCCTAGC 1300
 Db 2458 TAGTTTCTTGCTTCTCTAGGCTATAGCTTGCCTCTATGTTGGGCTTTTACCATTATTT 2399
 QY 1301 CTTCTT--TAGTCTGCAATTTGAAGATAGATATTTTACATCTCATCTTATTTATCTTGAAT 1358
 Db 2398 ATCTTTGTAGTGTGGAATTTGTGAAGAATATTTGTGAAATTTGTTTGTCTGATGGAAT 2339
 QY 1359 GTCTTTTCTTCTCAACTATTGTGACAGAAAGTTTTTCTAAGTGCAGTGTCTTGGCCCTGA 1418
 Db 2338 GTCTTTGTTTCTCCATCAATGTTAATGAGAGTTTGTGATACAGTAACCTGGGCTGG 2279
 QY 1419 CATCTGTAGTCTCTGGAGTCTGTAGCACATCTGTGAGGCGCTTCTTACATTTGAGTT 1478
 Db 2278 CATTTGTGTTCTCTTAGGCTCTGTATAACATCAGTCCAGGATCTTCTGGCCCTTCATAGT 2219
 QY 1479 TCTATTGAAAGAGTCAGGTGTAATTTCTAATACATCTGCTTTTATATGTTAATTTGCTCTT 1538
 Db 2218 TCT-GGCGAGAAGTCTGCTGTGATTTCTGATAGGCTCCCTTTATATGTTACTT-GACCTT 2161
 QY 1539 TTTCCCTTGCATCTTTAAATATCTTTCTTTGTTCTATACATTTTGTAGTCAATTTGATTA 1598
 Db 2160 TTTCCCTTACTGTTTAAATATCTTTCTTTTATTTTGTGCGTTTGGTGTGTTGCAATTA 2101
 QY 1599 TGCACGTGCGGAGTCTTTTCCGTCCTCAATCTATTTGGTGTGTTTGTATGCTTCTTGTGA 1658
 Db 2100 TGTGAGGAGAGTGTCTTTTCTGTTCCAACTATTTGGAGTCTGTAGGCTTCTTGTG 2041
 QY 1659 CTTGATAGGCACTCTTTTCTCAAGGTTAGGAATTTTCTTTTGTGTTTCTTGTGAAAA 1718
 Db 2040 TGTCTATGGGTATCTCTTTTGTAGGTAGGAAAGTTTCTTCTATGTTTGTGTTGGA 1981
 QY 1719 TATTTTCCCTCTTTTGACCTG-----CCTTCTCCCTCTCTCTATTCCT 1764
 Db 1980 TATTTACTTGGTCTTTTGTAGCTGGGAGTCTTCACTCTCTTCTATACCTATTTATCTTAGGT 1921
 QY 1765 TTGGTTTTTGATAGTGTCTCTGGCTTCTCTGGATGTTTATGCTCGGATTTATTTAGCT 1824

Qy	2004	CTTTGAGGTCCTGAAGAATGTTTACTCAATTTCCCTCCAGTAATTTACATTTTACATTTTCAATAGGTTT	2063
Db	1680	TTTTAAATTCCTCAACTGTTT-----GATTGTGTTTTCCTGGAATTC	1638
Qy	2064	CTTTAATGGAATTTATTCATTTTCCTCTCAAGGACCTTTTATGAATTCATAAAATCTATGT	2123
Db	1637	CTTTACAGGATTTTGTGTGTCCTCTCTATGGGCTTCTACTGTGTTTATTATGTTTCTCT	1578
Qy	2124	TAAGTCTCTT	2133
Db	1577	GGAAATCTTT	1568
RESULT 8			
AAx83007/c			
ID	AAx83007 standard; DNA; 51259 BP.		
XX	AAx83007;		
AC	XX		
KW	XX		
DT	31-AUG-1999 (first entry)		
XX	XX		
XX	Partial mouse WRN genomic sequence #3.		
XX	Mouse; WRN; Werner's syndrome; detection; diagnosis; autosomal;		
KW	recessive disorder; phenotype; ss.		
XX	XX		
OS	Mus musculus.		
XX	XX		
XX	WO9724435-A1.		
XX	XX		
PD	10-JUL-1997.		
XX	XX		
XX	30-DEC-1996; 96WO-US020785.		
PF	XX		
PR	28-DEC-1995; 95US-0009409P.		
PR	29-DEC-1995; 95US-00580539.		
PR	30-JAN-1996; 96US-0010835P.		
PR	30-JAN-1996; 96US-00594242.		
PR	12-APR-1996; 96US-00632175.		
XX	XX		
PA	(DARW-) DARWIN MOLECULAR CORP.		
XX	XX		
PI	Oshima J, Fu Y, Yu C, Mulligan J, Schellenberg GD;		
XX	XX		
DR	WPI; 1997-363671/33.		
XX	XX		
PT	Isolated nucleic acid molecule encoding the WRN gene product - useful for		
PT	detection and treatment of Werner's syndrome, and related diseases.		
XX	XX		
PS	Claim 1; Fig 7; 153pp; English.		
XX	XX		
CC	This sequence represents a fragment of the genomic sequence containing		
CC	the coding region for the mouse WRN gene (AAx83004). The corresponding		
CC	human gene (AAx83001) encodes a protein related to Werner's syndrome. The		
CC	products can be used for the detection and treatment of Werner's syndrome		
CC	(WS), an autosomal recessive disorder with a complex phenotype, as well		
CC	as related diseases		
XX	XX		
XX	XX		
SQ	Sequence 51259 BP; 14533 A; 9635 C; 10266 G; 16825 T; 0 U; 0 Other;		
Query Match 24.8%; Score 565.399; DB 1; Length 51259;			
Best Local Similarity 62.1%; Pred. No. 1.4;			
Matches 1155; Conservative 0; Mismatches 576; Indels 128; Gaps 18;			
Qy	292	TCTATTCTTGATTTCTATCTTTGGCTCATTTTTAACTCAGTAGTCAGTTGTTTGGTTTCC	351
Db	10412	TCTTATTTTCATTTCTGCTTGACCAAGGTATCATTCAGTAGCGCTGTTTCAGTTCC	10353
Qy	352	ATAGTTTCTAGTTTCTGTTTCTGTTGTTGTTGTTATCTAGATTTAAGCTGT	411
Db	10352	ATATGTATGTGCTTTCGGTTGTTGTTGTTGTTGTTTAAAGACCAATCTTAGCCGCTGT	10293
Qy	412	GGTGGTCAGATAGGACATAGAGTATTATTTCAATTGCTCTTTTATCTGTGCGACTTGCTT	471

Db 10292 GGTCTGATAGAGGTGCGATGGATGATTTCCATCATCTTTGAATCTGTAGAAGTCGTGTT 10233
Qy 472 TGTGTTGAAATATGTAATTAATTTGGAGA--GTTTCATAGGCTGTGACAGAAGGTAC 529
Db 10232 TGTGACCAGCATATGCTGATGTTTGGAGAAGGTTCCATAGGCTGTGAGAAGGTAT 10173
Qy 530 AGTCTTTGCTTTTGGTGAATAGTCTGTAATATCT-CTAGGTCACATGGTGTATGAC 588
Db 10172 ATTATTTGCTTTGGATGACATGTTCTATATAATATCTGTAGATCCATTTGGTTCATAAC 10113
Qy 589 ATCAGTTAGTCCAGCAATTTCTCTGTTTCGTTTTTGGTGGATGACCTAACCTGTTGGAG 648
Db 10112 ATCTGTTAGTTTCACTGTCTCTGATTTAGTTTCTGTTCCGTGATCTCTCCATGCTG 10053
Qy 649 AGAATGGGTATTGAAGTACCCACATCTGCTGTGAGGT-CAATATGTTGATTTAGCT 707
Db 10052 AGAGTGGGTGCTGAAATCTCCACATATATTGATATCGGGTGTGATGTGTCCTTACAT 9993
Qy 708 GTAGCTGTCTGTTGTTTATGAACCTGGGTGACATTTGTTTGGTGCATAGACATTAAGAA 767
Db 9992 TTAGTAAAGTTTCTTTATGAATGTGGTGCCTTGCATTTGTAGCATACATGTCAGAA 9933
Qy 768 TTGCAATGCTCTCTGTTGGAT-TTTCCTTTGATGCTTATGTAGTATTTCTCCCATCTC 826
Db 9932 TTGAGAGTTCACTGTGACAGATGTTCCCTTTGACCAGTATGAAGTGTCTCTTATCTT 9873
Qy 827 ATCTGCTTAGTTTGGGTTT-----AAGTCTATTAGTCAGATATAAATGACTGTATC 880
Db 9872 TTTTTTTTTTGTAACTTGGTTGAGAGTTGAAATTTATTCATATTAGAAATGCTACTOC 9813
Qy 881 GGCTGTGCTCTTAG-----GCCATTTGCTTAGAATA-TCCTTCCATCTCTTTTACTCTA 934
Db 9812 AGCTTGTCTTGGGAAACCAACATTTGCTTGGAAATTTGTTTCCAACTTGAACCTCTG 9753
Qy 935 AGTGATGCTATCCAGT---GTAGGTGTCTTTTGGATGACAGCAGTAGTAGGATGATCT 991
Db 9752 AGGTAGTGTCTGCTTGTCTGCTGAGGTGCAATTCCTGTATGACGAAATGCTGGTCC 9693
Qy 992 TGTGTTTCATATCCATCTGTTPACCAGTATCTTTTCTAGAGAAATTAAGATCATTTAGT 1051
Db 9692 TGCCTTACATCCAGTCTGTATGCTATGCTCTTTTGTAGGAATTCG-----GTC 9642
Qy 1052 CATTTAGTTTGAATATTAATGACAGCAGTGTGTGGAATCTTGTA-TCCTGCACTTG 1110
Db 9641 CATTTAGTTTGAAGATATTAAGGAAAGTATTTGTTACTCTCTGTTATTTTGTGTTG 9582
Qy 1111 TGAAGT 1170
Db 9581 TTAGAGTGGAAATATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9537
Qy 1171 GT 1230
Db 9536 ----- 9537
Qy 1231 CTGGAATTTATTTATTTATTTATTTCTTTGTAATGTGGTAACTCTTTAGATTTGAAGTTT 1290
Db 9536 -TTGAAAGATTACTTTCTGCTTCTTTCTTTAGGGTGTAGTTTCCCTCTGTGTGTGTGT 9478
Qy 1291 TTTCTCTTA--GCCTTCTTTAGTCTGCAATTTGAAGATAGATATTTCTTTTACATCTGATTT 1348
Db 9477 TCCATCTATTATCTTTTGTAGAGCTGATTTGTGGAATGATTTGTGTAATTTTGGTTTT 9418
Qy 1349 ATCTTAGAATGCTTTCTTTCTCCAACTATTGTGACAGAAATTTTCTTAAGTGCAGTAG 1408
Db 9417 GTCATGGAATATCTTGGTTTCTCCATCTATGATTAATTCAGAGTTTGTGGGTATAGTAG 9358
Qy 1409 TCTGCCCTGACATCTGTAGTCTCTTTGGAGTCTGTAGCACATCTGTGAGGGCTCTTCTAC 1468
Db 9357 CTTGGCTGGCAATTTGTGTTCTCTTTGGTCTGTATGACATCTGCCFAGGATCTCTAGC 9298
Qy 1469 ATTTTGAAGTTCTATTTGGAAAGTCAGGTGTAATTTCTTAATACATCTGCCCTTTATGTTA 1528

Db 9297 TTTCTAGTCTCTGTGTGAGAACTCTAGTCTCTATATCTGATAGGCTGCTTTATATGTTA 9238
Qy 1529 ATTGTCTCTTTTCCCTTGGCATCTTTTAATATTTCTTTCTTTCTTTTACTTTAGTGAT 1588
Db 9237 CTT-GACTTCTTTCCCTTACTGCTTTCATTAATCTTTCTTTTGTAAAGCAATTTGATGT 9179
Qy 1589 TTGATATTATGACACTGTGGGAGTTCCTTTCCGGTCCAAATCTATTTGGTGTGTTGAT 1648
Db 9178 TCAATTTATGTGACAGAGGA-ATTCTTTTCTTTGTCAGTCTATTTGGAGTCTCGAG 9120
Qy 1649 GCTTCTGTACCTGATAGGATCTCTTCTCAAGGTTAGAAATTTTCTTTTGTGTT 1708
Db 9119 GTTCTTGTATGTTCATGACATCTCTTTCTTTAGTTATGGAAGTTTCTTCTATAAT 9060
Qy 1709 TTTCTTGAATAATTTTCCCTGCTTTTGACCT-GCCTTCTTCCCTTCTCTATTC- 1763
Db 9059 TTGTTGAAGATATTTACTTGGCCCTTGTAGTTGGGAATCTTCACTTTCTTCTATACCTAT 9000
Qy 1764 -----TTTGGTTTGTGATAGTGTCTCTGGCTTCTGATCTTTTATGCCCTGGAT 1814
Db 8999 ATCTTAGGTTTCAATTTCTCAATGTGTCCTATATTTCTCTGATGTTTGGTTAGGAC 8940
Qy 1815 ATTTAGACCTTAACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTG 1874
Db 8939 TTTTTCATTTTGTATTTCTTTGACTGTGTGTCAATGTTTCTTCTATGTAATCTTCTGCA 8880
Qy 1875 CTGAGATTTCTCTCTATCTCTTGTATTTCTGTGCTGAGTGTGCTCTGAGGTTCTT 1934
Db 8879 CCTGAGATTTCTCTCTATCTCTTGTATTTCTGTTGGTGATCTGTATATATGACTGCT 8820
Qy 1935 G-TTGGTTTCTTAATTTTCTTCAATTTCCAGATTTCTTCAATTTGGTGTGTTTATTAAT 1993
Db 8819 GATCTTTTCTTAGGTTTCTGTCTCCAGCTGTGTCTCCCTTGTGATTTCTTTATGTT 8760
Qy 1994 TCTATTTCCACTTTCAGGTCCTGAAATGTTTACTCATTTTCTCTCC-AGTATTTACA 2050
Db 8759 TCTAGTTCCATTTTAGATCTCGGATGTTTGTTCATTTCTTCACTTCTACCTGTTGATTTG 8700
Qy 2051 TTTTCAATAGTTTCTTTTATGATTTATTCATTTCTCTCTTCAAGGACCTTTTATGAAT 2109
Db 8699 TTTCTCTCTAGTTCTTTAAAGATTTTGTGTTTCTTCTTAAAGGCTTCTAGTTGTTT 8641

RESULT 9

ADAO2627
ID ADAO2627 standard; DNA; 47115 Bp.

XX

AC ADA02627;

XX

AC ADA02627;

DT 06-NOV-2003 (first entry)

XX

DE Mouse Flt3 carcinoma associated gene, SEQ ID NO:1145.

XX

XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;

XX

XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;

XX

XX gene; ds.

XX

XX Mus sp.

XX

XX WO2003057146-A2.

XX

XX 17-JUL-2003.

XX

XX 26-DEC-2002; 2002WO-US041414.

XX

XX 26-DEC-2001; 2001US-00035832.

XX

XX (SAGR-) SAGRES DISCOVERY.

XX

XX Morris DW;

XX

XX WPI; 2003-587068/55.

XX

XX

PT New recombinant nucleic acid encoding carcinoma associated protein,
 PT useful for preparing compositions for treating carcinomas.

XX Claim 1; SEQ ID NO 1145; 245pp; English.

XX The invention relates to recombinant carcinoma associated (CA) nucleic
 CC acid sequences from mouse and human (ADA01482-ADA03094), and to
 CC recombinant carcinoma associated proteins (CAP) encoded by them. The
 CC invention also encompasses expression vectors and host cells comprising a
 CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
 CC binds to the protein, and a biochip comprising CA nucleic acid or
 CC fragments thereof. The sequences of the invention were identified using
 CC oncogenic retroviruses, which insert into the genome of the host organism
 CC at random. Many of these do not carry transduced host oncogenes or
 CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
 CC direct consequence of the effects of proviral integration into host
 CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
 CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
 CC leukaemia) or a propensity to carcinoma by determination of the sequence
 CC of a CA gene, or by determination of CA gene expression in particular
 CC tissues. CA nucleic acids, proteins and antibodies are also useful as
 CC therapeutic agents and in screening and evaluating drug candidates. The
 CC present sequence represents a specifically claimed murine CA nucleic acid
 CC sequence of the invention. Note: The complete sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 47115 BP; 11483 A; 10181 C; 10540 G; 13012 T; 0 U; 1899 Other;

Query Match 24.4%; Score 555.699; DB 1; Length 47115;
 Best Local Similarity 62.5%; Pred. No. 1.6;
 Matches 1159; Conservative 0; Mismatches 563; Indels 131; Gaps 20;

292 TCTATTTCTGATTTCTATCTGCTCATTTTAACTCAGTAGGAGCTTTGTTGTTCC 351
 30363 TTTCTTTTATTTCTTCTTGACCAAGTATCATTCAGCAGAGCATTTGTCAGCTTCT 30422

352 ATAAAGTTGTAAGTTTCTGTTGTTCTGTTGTTGTTGTTATCTAGATTTAAGCTGT 411
 30423 GTGTGTATGTGCTTCTGTTGTTT---TTGTTGGTATTAAAGCAGGCTTAGTCTGT 30479

412 GGTGTCAGATAGACATAGATATTTTCAATTGCTTTTATCTGTCAGACTTGCCT 471
 30480 GGTGATCTGATAGGATATAGGGAATTTTCAAGTCTTCTGTTATCTGTTGAGGCTGTTT 30539

472 TGTTTTGAATATATATTCAATTTTGAGA--GTTTCATAGGGTGTGCAAGAAGGTAC 529
 30540 TGTGACCAATTATATGTCAGTTTGAAGAAGTACTGTGAGGTCTGAGAAGAGGTAT 30599

530 AGTCTTTGTTGTTGG--TGAAATAGTCTGTAATATCT-CTAGGTCACTGTTGTTATGA 587
 30600 ATTCTTTGCTTTAGGATGAATGTTCTATAAATACCTGTTAGATCCATTTGCTTCATAA 30659

588 CATCAGTTAGTCTCAGCATTTCTCTGTTTCTGTTTGTGTTGATGACCTAACTGTTGGA 647
 30660 TTTCGTGTTAGTTTCACTGTGTCGATTTAG--TTTCTGTTTCCATGATCTGTCCTAATGCT 30718

648 GAGAACTGGGTATGAAGTAGCCACATCTGTTGTGTGAGGT-CAATATGTGATTTTAGC 706
 30719 GAGAGTGGGTGTTGAAGTCTCCCGCTATTATTGTTGTGGGTGCAATGTGTGCTTTGAGC 30778

707 TGTAGCTGCTGTTTGTATGAATCTGGGTGACATTTGTTGTCATAGACATTAAGA 766
 30779 TTTAGTAAAGTCTCTTTTATGACTGTGGATGCCCTTGCAATTTGGAGCATAGATGTTTCA 30838

767 ATTGCAATGTCCTCTTGGTGA--TTTCTCTTTGATGCTATGTAGTATCTTCCCAATCT 825
 30839 ACTGAGAAATCACTCTTAGATTTTCTCTTTCAGCAGTATAAAGTGTCTCTCTTATCT 30898

826 CATCTGCTTAGTTTGGTTTAGTCTA--TTAGTCAGATATTAAGTACTGTATCGGCT 884
 30899 TTTTGTATACTTTTGGTTGAAGAGTTGATTTTATTCGATATTAGAATGATATCCAGCT 30958

QY 885 TGCCTTCTTAGGGCCATTTCCTTTAGAAATA--TCTTTTCCATCCCTTTTACTCTAAGGTGATGT 943
 DB 30959 TGTTCCTGAAAAAATTTGCTTTGGAAAAATGTTTTCCAGACTTTTACTCTGAGGTCGTGT 31018

QY 944 CTATCCATG--GTAGGTTGTCTTTTGGATGACAGCAGTAGATGGATCTTGTGTTTCAT 1000
 DB 31019 CAGTCTTGTCTACTGAGGTGCAATTTCCGTGTGCGAGCAAAATGCTGGGTCTCTGTTATGT 31078

QY 1001 ATCCATTTCTGTACCCAGPATCTTTTCTPAGAGAAATTAAGATCAATGAGTCATTGATGT 1060
 DB 31079 ATCCAGTCTGTAGTCTATGTCTTTTATTGGGAATTGAG-----TCCATTGATGT 31130

QY 1061 TGAGATTAATCAATGAGCAGTGTGTTGTGATTCCTTTATCTTGCACATGTGGAAGTGT 1120
 DB 31131 TAATAGATTAATAGGAAAAAGTGAATGTTACTTCTCTGTTATTTTGTGTTAGAGGTGAA 31190

QY 1121 GT 1180
 DB 31191 TTATGTTTGTGTGGCTATCTTCTCTCTGGTGTGT----- 31224

QY 1181 GT 1240
 DB 31225 ----- 31233

QY 1241 TTATTAATTCATATTTTCTGAAATGTTGGGTAAACATCTTTAGATTTGAAGTTTTCCTCCTAGC 1300
 DB 31234 ACTTCTCTGTTTCTAGGATGTAGTTTCTCTCTCTGTAAGTGTGTTTCCATCTATAT 31293

QY 1301 CTCTCTTAGGTCTGATTTGAAGATAGATATCTTACATCTGATTTTATCTTGAATGT 1360
 DB 31294 CCTTTGTAGGGCTGATTTTATGGAAGATACCTGTGTAATTTGGTTTTGTCTATGGAATAT 31353

QY 1361 CTCTCTCTCCAACTATTGTGACAGAAAGTTTCTAAGTGCAGTAGCTGCGGCTGACA 1420
 DB 31354 CTTGGTTTCTCTGCTCTAGTGGCAATTTAGAGTTTTGCTGGGTATATAGTCTGCGGCTGGCA 31413

QY 1421 TCTGTAGTCTCTGTGAGTCTGTAGCACATCTGTGCGAGGCTTCTTACATTTTGTGTTTC 1480
 DB 31414 TTTGTGTTCTCTTAGGGTCTGTATGACATCTGCCAGGATCTTCTGCTTTCATAGTCTC 31473

QY 1481 TATTCGAAAGTCAGGTCTTAATCTTACATCTGCTTTATATGTTAATTTGCTCTTTT 1540
 DB 31474 TGGT-GAGAGTCTGGGTGTAATCTGTAGTGTCTCTTATATGTTACTT-GACCTTTT 31531

QY 1541 TCCCTTGATCTTTTAAATATCTTCTTGTCTATCTACTTTTAGTGTGATTTGATTTATATG 1600
 DB 31532 TCCCTTACTGCTTTTAAATATTTATTTCCITTTGTAGTGCATTTGGTGTGTTGATTTATG 31591

QY 1601 CACTGTGGGAGTTTCTTTTCCGGTCCAATCTATTTGGTGTGTTGTATGCTTCTGTGACC 1660
 DB 31592 TGATGGGAGGAATTTCTCTCTGTCGAGTCTATTTGGAGTCTGTAGGCTTCTGTGATG 31651

QY 1661 TTGATGAGCATCTCTTTCTCAAGGTGTTAGGAAATTTTCTTTTGTGTTTCTTCAAAATA 1720
 DB 31652 TTCAAGGCACTCTTCTTTTAGGTAGGAGAAATTTCTCTCTATAAATTTGTTGTAAGATA 31711

QY 1721 TTTTCCCTGCTTTTGACCTGCTCTTCTTCCCTTCC-----TCTATCTCTTT 1766
 DB 31712 TTTACTGGCCCTTTAAAGTTGGGAATCTTCACTCTCTCTATAACTACTATCTCTAGGTTT 31771

QY 1767 GGTTTTTCATAGTGTCTCTGGCTTCTGGATGTTTTTATGCTGCAATTTATTTAGACTTA 1826
 DB 31772 GGTCTTCTCATGTATCTCGAATTTCTTGGATGTTTTGGGTAGGAGCTTTTGTGCTTTT 31831

QY 1827 A-CATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCTGCTGCTGAGATCT 1885
 DB 31832 TGGTTTTCTTTAACTGCAAGTGCAATGTTTCTTCTATGATTTTCTATACCTAAGATTC- 31890

QY 1886 CTCCTCTATCTCTGTTATCTGTGAGGCTTGTCTGAGGTTCTCTGAGTTTCTGTTGGTT-CT 1944
 DB 31891 -----TCTCTGTATTTCTGTTGGTGTGATGCTTGCATCTATGACTCTCTCTCCC 31942

Db 31474 TGGT-GAGAGTCGGTGAATTCGATAGTCTGCTTTATATGTTACTT-GACCTTTT 31531
Qy 1541 TCCCTTGGCATCTTTTAATATCTTTCTTGTGTTCTATATCTTTTAGTGAATTAATATG 1600
Db 31532 TCCCTTACTGCTTTTAATATATCTTTCTTGTGTTCTATATCTTTTAGTGAATTAATATG 31591
Qy 1601 CACTGTGGGAGTCTTTTCCGGTCCCAATATTTTGGTGTGTTGATGCTCTTGAC 1660
Db 31592 TGATGGGAGGAATTTCTTCTCTGGTGGAGCTATTTGGAGTCTCTGAGCTCTTGATG 31651
Qy 1661 TTGATPAGGCATCTCTTTCTCAAGGTTAGGAAATTTCTTTTGGTGTGTTTCTTGAATA 1720
Db 31652 TTCAATGGGCACTCTTTCTTTAGGTTAGGAAATTTCTTCTATAATTTGTTAAAGATA 31711
Qy 1721 TTTTCCCTGCTTTGACCTGCTTCTTCCCTTCC-----TCTATTCCTTT 1766
Db 31712 TTTACTGGCCCTTTAAGTTGGGAATTTCACTCTCTCTATAAATACTATCTTAGGTTT 31771
Qy 1767 GGTTTTTCATAGTCTCTCGCTTCCCTGGATGTTTATGCTCGAATATTTAGACTTA 1826
Db 31772 GGTCTTCTCATGTATCTCGATTTCCGGAATTTTGGGTTAGGAGCTTTTGTCTTT 31831
Qy 1827 A-CATTTCTTTGACCAAGGTAATCAATTTCTTATCTTGTCTTCACTGCTGAGATCT 1885
Db 31832 TCCGTTTCTTTAACTGAGTCCCAATGTTTCTATGATATTTCTATACCTAAGATTC- 31890
Qy 1886 CTCCTTCTATCTCTGATCTTCTGAGTGGGCTGCTCTGAGGTTCTGTTGGTT-CT 1944
Db 31891 -----TCTCTGTATCTTGGTGATCTGTCATCTATGACTCTGATCTCTCCC 31942
Qy 1945 TAAATTTTTCATTTCCAGATTTCTTCAAGTTGGGTTTGTATTAATTTATTTCCAC 2004
Db 31943 TAGGTTTCTATCTCCGGTGTGCTCCCTTTGTGATTTCTTTATTTGTTCTATTTCCA- 32001
Qy 2005 TTTGAGTCTTGAATGTTTACTCAATTTTCTCC---AGTATTTACATTTTCATAGGT 2061
Db 32002 TTTTGAATCTTGGATAGTTTGTGTCCTTTCAATTTCTTCCCTGTTGATGTTTCTCTGATA 32061
Qy 2062 TTTTGAATCTTGGATAGTTTGTGTCCTTTCAAGGACCTTTTATGAATTCATA 2114
Db 32062 TTTTGAATCTTGGATAGTTTGTGTCCTTTCAAGGACCTTTTATGAATTCATA 32114

RESULT 11
ID ADA02501
XX ADA02501 standard; DNA; 96597 BP.
AC ADA02501;
XX
DT 06-NOV-2003 (first entry)
DE Mouse Bach2 carcinoma associated gene, SEQ ID NO:1019.
KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX
OS Mus sp.
XX
PN WO2003057146-A2.
PD 17-JUL-2003.
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
PR 26-DEC-2001; 2001US-00035832.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-587068/55.

XX New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX
PS Claim 1; SEQ ID NO 1019; 245pp; English.
XX
CC The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed murine CA nucleic acid
CC sequence of the invention. Note: the complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 96597 BP; 25792 A; 20376 C; 20447 G; 28650 T; 0 U; 1332 Other;

Query Match 23.1%; Score 525.898; DB 1; Length 96597;
Best Local Similarity 61.9%; Pred. No. 1.6; Mismatches 556; Indels 147; Gaps 21;
Matches 1143; Conservative 0;
Qy 285 AAACACTCTTATTTCTGATTTCTATCTGCTGCTATTTTAACTCAGTAGTGAATTTT 344
Db 51129 AAAGTTTAAATTTCTTATTTCTTCAATGACCAATATCATTCATTGAGGCTTTTC 51188
Qy 345 GGTTCCTCATPAAGTTGTAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 404
Db 51189 AGCTTCATGCTATGCTGCTTTCCCTGTT-----TTCTGTTATTTGAAGACCGCTTT 51244
Qy 405 AGCTGCTGCTGCATAGGACATAGATATTTTCAATTTCTTTTATCTGTCGAGA 464
Db 51245 ACTCCATGGTGAATCATAGGATGATGGGATATTTCAATCTTCTGTTCTGTTGAGG 51304
Qy 465 CTTCCTTTCTTTTGAATATGATTTCAATTTTGGAGA--GTTTCATAGGGTCTGCACAG 522
Db 51305 CCTGTTTCTGAACCAATATATGTTGTCAGTTTTCAGAGGTACCATGAGTCTCGAAG 51364
Qy 523 AAGGTACAGTCTTTGTTT---GGTGAATAGTCTGTAAATA--TCTCTAGGTCCATT 578
Db 51365 AAGGTACATTTTCTTTTCTTTAGGATGAATGTTCTATAAATAATCTGCTAAACCTTTA 51424
Qy 579 GGTATTATGACATCAGTTAGTCTCCAGCATTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 638
Db 51425 GGTCTTAACTTCTATTAGTTTCACTGCTCTCTGTTAG-TTTCTGTTTCTATAATCTG 51483
Qy 639 ACTGTTGGAGAGATGGGGTATTGAAGTAGCCCATCTGTTGTTGTTGTTGTTGTTGTTGTT 697
Db 51484 TCCATTGCTGAGAGTGGGGTGTGAAGTCTCTACTATTATTGTTGGGTGCAATGTT 51543
Qy 698 GATTTTGTAGCTGTGCTGTTGTTTATGAATCTGGGTGACATTTGTTTGGTGCATAG 757
Db 51544 GCTTTGAGCT-----TAAGTTTCTTATGCAATGTTGTTTCTTGCATTGGAGACAG 51598
Qy 758 ACATTAAGAATTCATGCTCTTGGTGA-TTTTCCCTTTGATGCCCTATGATGATTTCT 816
Db 51599 ATCTTCAGAAATGAGAGTTCAATTTGGTAGATTTTTCCTTTGACTAGTATGAAGTCT 51658
Qy 817 TCCCAATCTCATCTGTTAGTTTGGGTTTAAAG-TCTATTAGTCAGATTTAAATGACT 875

XX Oshima J, Fu Y, Yu C, Mulligan J, Schellenberg GD;
PI WPI; 1997-363671/33.
XX Isolated nucleic acid molecule encoding the WRN gene product - useful for
DR detection and treatment of Werner's syndrome, and related diseases.
XX Claim 1; Fig 7; 153pp; English.
XX This sequence represents a fragment of the genomic sequence containing
CC the coding region for the mouse WRN gene (AA83004). The corresponding
CC human gene (AA83001) encodes a protein related to Werner's syndrome. The
CC products can be used for the detection and treatment of Werner's syndrome
CC (WS), an autosomal recessive disorder with a complex phenotype, as well
CC as related diseases
XX
SQ Sequence 51259 BP; 14533 A; 9635 C; 10266 G; 16825 T; 0 U; 0 Other;
Query Match 22.9%; Score 521.797; DB 1; Length 51259;
Best Local Similarity 60.7%; Pred. No. 2.2;
Matches 1118; Conservative 0; Mismatches 637; Indels 88; Gaps 17;
QY 232 TCTATTTCTGATTTCTATCTGCTCATTTTAACTCAGTAGAGTGTGTTGCTCC 351
DB 26099 TCTTCATTTTCATTTCTGCTGACCAAGTTATCAATGAGTAGAGCGCTGTCACTCC 26158
QY 352 ATAGTTTGTAAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTATCTAGATTAAAGCTGT 411
DB 26159 ATATGATGTGCTTCCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 26218
QY 412 GGTGTCAGATAGACATAGAGTATTTTCAATGCTTTTATCTGTCGAGACTTGCTT 471
DB 26219 GGTCTGATGATAGGTCGATGGAGTATTCATCATCTGCAATCTGTAGAACTGCTTT 26278
QY 472 TGTTTTGAATATGATTTCAATTTGGAGA--GTTTCATAGGTCGTCACAAAGATAC 529
DB 26279 TGTGACAGCTATATGTCAGTTTGGAGAGGTTCCATGAGGTCGTGAGAAAGGTAT 26338
QY 530 AGTCTTTGTTTGGTGAATAGTCTGTAAATATCT-CTAGGTCACCTTGTTTATGAC 588
DB 26339 ATTTTTCCTTTGGATGACATGTTCTATAAATATCTGTAGATCCATTTGGTTTCAATC 26398
QY 589 ATCAGTAGCTCCAGCATTTCTGTTTCGTTTGTGTTGTTGATGACCTAACTGTTGAG 648
DB 26399 ATCTGTTAGTTTCACTGTCGTCGCTTAGTTTCTGTTTCCGTCATCTGTCCTGCTG 26458
QY 649 AGAATGGGTATTTGAAGTAGCCCACTATCTGTGTGAGGTCAATATGATTTAGCTG 708
DB 26459 AGAGTGGGTGCTGAAATCTCCCACTATTTGATCAGGTATGATGTGCTTTGAGAT 26518
QY 709 TAGCTGCTGTTTATGAACTTGGGTGAATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 768
DB 26519 TTAGTAAAGTTTATGAAATGTTGGTGGCTTGCATTTGGAGCATACTGTTTCAAGAT 26578
QY 769 TGCATGCTCTCTGCTGAT-TTTTCTTTGATGCTATGATGATGATTTCTTCCCAATCTCA 827
DB 26579 TGAGAGTTCACTTGGCAGATTTTCTTTGACCAATGAAAGTGTCTTCTTATCTTT 26638
QY 828 TCTGCTAGTTTGGGTTTAACTCT-ATTAGTCAGATATTAATAAGTGTATCGGCTTG 886
DB 26639 TTTTGTAACTTGGTTGAGAGTTGAAATTTATTCATATTAAGTGTGCTACTCCAGCTTG 26698
QY 887 CTCTCTTAGG-----GCCATTTGCTAGAAATA-TCTTTTCCATCTCTTTTACTCTAAGTGA 940
DB 26699 TTTCTGGGAACCAACCAATTTGCTTGGAAATTTGTTTCCAACTTGAATCTGAGTAG 26758
QY 941 TGTCTATCCATG---GTAGGTTGTTTCTTTTGGATGAGCAGTAGGATGGAATCTGTTTT 997
DB 26759 TGTCTGCTTGTGTCATGAGGTGCAATCTCTGATGAGCAAAATGCTGGGCTCTGTTA 26818
QY 998 CATATCCATCTGTTACCAGATCTTTTTCTAGAGAAATTAAGATCATGAGTCATGA 1057
|||||

DB 26819 CACACCAGTCTGTTAGTCTATGTTTCTTTTGGAGAAATGA-----GTCCATTGA 26869
QY 1058 TGTGAGAAATATCAATGAGCAGTGTGTTGGAATCTTGTGATCTTGTGCACTGTGGAAGTG 1117
DB 26870 TGTAAAGAGATATAAGGAAAGTGTGTTACTTCTCTGTTATTTT-----TGTTG 26920
QY 1118 TGTGCTG 1177
DB 26921 TTGTTAGAGGTGGAATTAAGTTTGTGTGGCTATCTCTTTTGGTTTGTGAAGATTGC 26980
QY 1178 TCTG 1237
DB 26981 TTTCTTGTCTTTTCTAGGGTGTAGTTTCCCTC-----CTTGTGTGTGTGTGTGTGTGTG 27025
QY 1238 TATTTATTAATCATATTTCTTGAATGTGGAACATCTTAGATTAAGATTTTCTCCT 1297
DB 27026 TTTCCATCTATTAATCTTTTGAAGCTGGAAGATATGTTGTAATTTGGTTTGTCAAG 27085
QY 1298 AGCCTTTCTTAGGTCTGCATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAA 1357
DB 27086 AAATACCTAGCAGCTTGACAGCACCTGAACACTCTAGAACTAAAGAAAGCAATACAC 27145
QY 1358 TGTCTTTCTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTCAGTAGTCTGCCCTG 1417
DB 27146 CCAAGAGGAGTAGACTGAGATTGGGAGTT-----TTGCCCTGGGCTG 27186
QY 1418 ACATCTGTAGTCTTCTTGGAGTCTGTAGCACATCTGTGCGAGGCCCTTCTTACATTTTGA 1477
DB 27187 GCATTTGTGTCTCTTAGGCTCTGATGACATCTGCTAGGATCTTTTAGCTTTCATAGT 27246
QY 1478 TTCTATTGGAAGAGTCAGGTGAATTTCTAAATACATCTGCTTTTATATGTTTAAATGTGCT 1537
DB 27247 TTTCTGTG-GAGAACTGTGTGAATTTCTGATAGGCTGCTTTATATGTACTTGACCT- 27304
QY 1538 TTTTCCCTTGTGATCTTTTAAATATCTTCTTGTGTTCTATCTTTTATAGTGAATTTAT 1597
DB 27305 TTTCCATTTGCTCTTTTAAATATCTTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTG 27363
QY 1598 ATGCACTGTGGGAGTTCTTTTCCGTTCCAACTATATTTGGTGTGTTTGTATGCTTCTGT 1657
DB 27364 ATGTGACAGGAGGAAATTTCTTTTCTGTCAGTCTATTTGGAGTTCTGGAGGCTTCTGC 27423
QY 1658 ACCTTATAGGATCTCTTCTTCAAGGTTAGGAATTTTCTTTTGGTGTGTTTCTTCTTGA 1717
DB 27424 ATGTTATGGGATCTGCTTTTATAGTTAGGAAGTTTCTTCTTATAAATTTGTTGAAG 27483
QY 1718 ATATTTTCCCTGCTTTTGTGACCTGCTTCTTCCCTTCTCTATTC----- 1763
DB 27484 ATATTTTACGGCTTTGAGTTGGGAATCTTCACTCTCTTCTATACATATATCTTAGG 27543
QY 1764 TTTGGTTTTGCAATGATGCTCTGCTGCTGCTGATGTTTATGCTGGATTTATTTAGAC 1823
DB 27544 TTTGGTCTTCTCATTTGCTGCTGGAATTTCTGATGTTTGGGTTAGGAGCTTTTGTGCAT 27603
QY 1824 TTAACATTTTCTTTGACCAAGGTATCCATTTCTCTATCTTGTCTTCACTGCTGAGATT 1883
DB 27604 TTTGTAATTTCTTTGATGCTGTTGTGCAATATTTTCTATGATCTTCTGCACTGAGATT 27663
QY 1884 CTCTCTTCTATCTTGTATCTG-TCAGTAGGCTGTCTCTGAGGTTTCTG-TTGGGT 1941
DB 27664 CTCTCTTCTATCTTGTATCTGTTTGGTGTGATGCTGCACTCTGCACTCTGATCTCTT 27723
QY 1942 TCTTAATTTTCAATTTCCAGATTTTCTTCACTGTTGGGTTTGTGTTTAAATCTTATTTTC 2001
DB 27724 TCTAGATTTTCTAACTCCAGGTTGTCTCCCTTTGTGATTTCTTTATTTGTTTCTAGTTTC 27783
QY 2002 CACTTTCAGGTCTCGAAATGTTTTTACTCATTTTCTCCCTCCAGTATTTTACA---TTTTCA 2058
DB 27784 CATTTTAGACTCTGGATGTTTGTTCATTTCTTCTGCTGTTTAAAGTGTGTTTCTGCG 27843
QY 2059 GGTTCCTTTAAGGATTTTATCATTTCTCTTCAAGGACCTTT 2101
DB 27844 TAATCTGTAGGAATTTTGTGTTTCTCTTTAAGGGCTTCT 27886

RESULT 14
 AAD51996
 ID AAD51996 standard; DNA; 9330 BP.
 XX
 AC AAD51996;
 XX
 DT 02-MAY-2003 (first entry)
 XX
 DE Mouse Cyp3A11 gene.
 XX
 KW Transcription control element; transgenic; transgenic animal; toxicity;
 KW drug metabolism; mouse; Cyp3A11; ds.
 XX
 OS Mus sp.
 XX
 XX Location/Qualifiers
 FH Key 5104..6218
 FT misc_feature /tag= a
 FT /note= "Non repeat region"
 FT misc_feature 6792..9330
 FT /tag= b
 FT /note= "Non repeat region"
 XX
 XX WO200288305-A2.
 PN
 PD 07-NOV-2002.
 XX
 XX 11-APR-2002; 2002WO-US011770.
 PF
 XX
 PR 12-APR-2001; 2001US-0283534P.
 XX
 XX (XENO-) XENOGEN CORP.
 PA
 XX Zhang W, Contrag P, Purchio A, Hashima S, Ma S, Nawotka K;
 PI
 XX WPI; 2003-120468/11.
 DR
 XX
 XX Identifying analytes that modulate expression of a reporter sequence
 PT mediated by human or mouse transcription control elements in transgenic
 PT rodent, comprises administering the analyte and monitoring expression of
 PT coding sequence.
 XX
 PS Claim 18; Fig 1B; 136pp; English.
 XX
 CC The invention relates to a method of identifying an analyte that
 CC modulates expression of a reporter sequence mediated by transcription
 CC control elements derived from a human Cyp3A4 gene, in a transgenic,
 CC living rodent. The method involves administering the analyte to the
 CC rodent and monitoring expression of the sequence of interest where an
 CC effect on expression indicates that the analyte affects the expression of
 CC the reporter sequence. The method is useful for identifying an analyte
 CC that modulates expression of a reporter sequence and in isolating and
 CC identifying mouse and human transcriptional control elements associated
 CC with cytochrome expression e.g. Cyp3A11 and CYP3A4 respectively. The
 CC cells and transgenic animals containing the expression cassette may be
 CC used as models for studies involving toxicity and drug metabolism. The
 CC present sequence is mouse Cyp3A11 gene
 XX
 SQ Sequence 9330 BP; 2360 A; 1897 C; 2014 G; 3056 T; 0 U; 3 Other;
 Query Match 21.3%; Score 486.098; DB 1; Length 9330;
 Best Local Similarity 61.9%; Pred. No. 6.4;
 Matches 1022; Conservative 0; Mismatches 509; Indels 121; Gaps 17;
 416 GTCAGATAGGACATAGAGTATTATTCATTGCTCTTTTATCTGTCGAGACTTGCTTTGTT 475
 2503 GTCGTAGTAGGATGGACAAATTCATATTTTGTATCTGTTGAGGCTGTTTGTG 2562
 476 TTGAATATGATTCATTTTGGAGA--GTTTCATAGGCTGCTGCACAGAGGTACAGTC 533
 2563 ACCAATTATATGTTAAATTTTGGAGAAGGTTCCGTTGGGTGCTGAGAAGTATAT---CAT 2619

QY 534 TTGTGTTTTTGGTGAATAGTCTGTAAATATCTCT-AGTCCACCTGGTTGTTATGACATCA 592
 DB 2620 TTGTGTTTTTAGGATAAAATGTTCTGTAGATACTGTCAAAATCCATTTGTTTCACTCTCT 2679
 QY 593 GTTAGCTCCAGCATTTCTCTGTTTTCGTTTTTGTGAGATGACCTAACTGTTGGAGAGAA 652
 DB 2680 GTTAGTTTCA--CTGTGCTCTGTTTAGTTTCTGTTTTCATGATCTGTCCACTGATGAAG 2737
 QY 653 TGGGGTATTGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGATGATTTTAGCTGTAG 711
 DB 2738 TGGTGTGTTGAAAGTCTCCCACTATTATTGTGTGAGGTGCAATGTGTGCTTTGAGCTTTAC 2797
 QY 712 CTGTGCTGTTTTATGAACCTTGGGTGACATTTGTTTGTGTCATAGACATTAAGAATTGC 771
 DB 2798 TAAAGTGTCTTTAATGAATGTGCTGCCCTTCAATTGGAGCATAGATATTCAAAATTGA 2857
 QY 772 AATGTCCTCTGGTGAATTTT-CCTTTGAAGCTATGTAGTATTCTTCCCAATCTCATCT 830
 DB 2858 GAGTTCTCTCTGGAGGATTTTACCTTTGATGAGTATGAAGTCTCCCTCTCTGTTCTTTTT 2917
 QY 831 GCTTAGTTTTTGGGTTTTAAGTCTA-TTAGTCAAGATATTAAATGACTGTATCGGCTTGCTT 889
 DB 2918 GATACTTTGGTTTGGAGTTGATTTATTGATATTAGAATGGCTACCCAGCTTGTT 2977
 QY 890 CTTAGGGCCATTTGCTTAGAATATC-TTTTCCATCTTTACTCTAAGTGAATGCTATC 948
 DB 2978 CTTAGACCATTTGCTTGGAAAATGTTTTCCAGGCTTTCACTCTGAGGTAGTGTCTGTC 3037
 QY 949 CAT--GCTAGGTGCTTTTTTGGATGCAGCAGTAGGATGGATCTGTTTTCATATCCA 1005
 DB 3038 TTTTCCCTGAGATGGTTTCTGTAAAGCAGCAGAAATGTTGGTCTCTGTTGTGTAGCCA 3097
 QY 1006 TTCTGTTACCCAGTATCTTTTTTCTAGAGAAATAAGATCAATTGAGTCATTGATGTTGAGA 1065
 DB 3098 GTCGTTAGTCTATGCTTTTTTATTGGGAAATTGAG-----TCCATTGATATAAGA 3149
 QY 1066 ATTATCAATGAGCAGTGTGTTGGGATCTTGTATCTTGCACCTTGTGAAGTGTGTGTG 1125
 DB 3150 GATATTAAAGAAAGTAATTTGTTGCTTCTATATTATTTTGTGTAGAGTTGGCAATCG 3209
 QY 1126 TGTGCTGT 1185
 DB 3210 TTTCTTTTGGCTGCTCTCTTTTGGCTTG----- 3237
 QY 1186 TGTGCTGT 1245
 DB 3238 -----TTGAGGAATTACTTT 3251
 QY 1246 ATTATATTTTTCTTGAATCTGGGTAAACATCTTTTAGATTGAACTTTTCTCTAGGCTTCT 1305
 DB 3252 TCTTGTCTTTTCTAGGGCGTGATATCTGCTGTATTTTTTTTCTGTATATCTTT 3311
 QY 1306 TTAGGTCGCAATTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAATGCTTTTC 1365
 DB 3312 GAAGGGCTGGA-TTCTGGAAGAGATAATGTGAATTTGGTTTTTGTGTCATGGAATCTTTGG 3370
 QY 1366 TTTCTCAACTATTGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTGAC--ATCT 1423
 DB 3371 TTTCTCCATCTATGGTAAATTGAGAGTTTGGCGGGTATAGTAGCCCTGGGCTGGCTTTTTT 3430
 QY 1424 GTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCCCTTCTTACATTTTGTGTTTCTAT 1483
 DB 3431 TTGTTCTCTTAGTGTCTGTATTAACATCTGTCCAGGCTCTTCTGGCTTTTCTAGTCTCTGG 3490
 QY 1484 TGAAGAAGTCAGGTGAATCTAAATACATCTGCTCTTATATGTTAACTGCTTTTCTTCC 1543
 DB 3491 T-GAAAAGTCGTGTGTAATTCGATAGGCTGCCCTTATATGTTACTT-GACCTTTCTCC 3548
 QY 1544 CTTGCACTTTTTTAATATCTTTCTTTTGTGTCTATCTATCTTTTAGTCATTTGATTTATGCA 1603
 DB 3549 CGTACTGCTTTTAATATCTCTTTTATTAGTGCATTTGTTGTTCTGATTTATTTGTGTGT 3608

Qy 1604 TGTGGGAGTTCCTTTCCGGTCCAAATCTATTTGGGCTTTTGTATGCTTCTTGACCTTG 1663
Db TGGGAGGAATCTCTTTTCTGTGTCAGTCTATATGAGGTTCTGTAGGCTTCTGTATGTTT 3668
Qy 1664 ATAGGCATCTTCTCAAGGTAGGAAATTTTCTTTTGTGTTTCTTCTTGAATAATTT 1723
Db ATGGGCATGTCATCTTTAGTTCGGGAAGTTTCTTCTATAATTTTGTGGAATAATTT 3728
Qy 1724 TCCCTGCTTTTGACCTG-----CCTTCTCCCTTCCCTCTATTCTTTGGT 1769
Db GCTGGCCCTTTAAGTTGAAATCTTCATCTCTCACTCACTTATTAATCTGTAGTTGGT 3788
Qy 1770 TTTTGCATAGTCTCTGGCTTCTGATGTTTATCCCTGGATATTTAGACTTAACA 1829
Db CTCTCATTTGTCCTGGATTTCTCTGATGTTTGAATAGGACCTTTTGTGTTTGA 3848
Qy 1830 TTTTCTTTTGACCAAGGTATCCATTTCTCTATCTTGTCTTCACTGCTGAGATCTCTCT 1889
Db TTATCTTTGATGTTGCTCGATGTTCTCTATGGAATCTTCTGCACCTGAGATCTCTCT 3908
Qy 1890 TCTATCTCTGTATCTGTGATGAGGCTGTCTCTGAGGTTCTG-TTGGGTTCTTAAT 1948
Db TCCATCTTTTGTATCTCTGTGCTGATGCTCAGCTATGTTCCAGATTTCTTCTCTAGA 3968
Qy 1949 TTTTTCATTTCCAGATTTCTCTCAGTTTGGGTTTGTATTAATCTTATTCACCTTTC 2008
Db GTTCTATCTCCAGGTTGCTCTGTTGGGTTTCTTATGTTGCTACTTCCCTTTT 4028
Qy 2009 AGGTCCTGAAATGTTTACTCATTTTCTCTCC 2040
Db AGGTCAGTAGGCTTTGTTTCATTTCCATCAC 4060

RESULT 15

AA51995
ID AAD51995 standard; DNA; 12275 BP.
XX AC AAD51995;
XX DX 02-MAY-2003 (first entry)
XX DE Mouse Cyp3A11 promoter.
XX KW Transcription control element; transgenic; transgenic animal; toxicity;
XX KW drug metabolism; mouse; promoter; Cyp3A11; ds.
XX OS Mus sp.
FH Key Location/Qualifiers
FT misc_signal 1..11002
FT /tag= a
FT /note= "Transcriptional control element"
FT TATA_signal 10884..10887
FT /tag= b
FT misc_signal 10914
FT /tag= c
FT /note= "Transcription start site"
FT misc_signal 11003..11005
FT /tag= d
FT /note= "Translation initiation codon"

WO200288305-A2.

07-NOV-2002.

11-APR-2002; 2002WO-US011770.

12-APR-2001; 2001US-0283534P.

(XENO-) XENOGEN CORP.

Zhang W, Contag P, Purchio A, Hashima S, Ma S, Nawotka K;

DR WPI; 2003-120468/11.
XX Identifying analytes that modulate expression of a reporter sequence
PT mediated by human or mouse transcription control elements in transgenic
PT rodent, comprises administering the analyte and monitoring expression of
PT coding sequence.
XX Claim 17; Fig 1A; 136pp; English.
XX The invention relates to a method of identifying an analyte that
CC modulates expression of a reporter sequence mediated by transcription
CC control elements derived from a human Cyp3A4 gene, in a transgenic,
CC living rodent. The method involves administering the analyte to the
CC rodent and monitoring expression of the sequence of interest where an
CC effect on expression indicates that the analyte affects the expression of
CC the reporter sequence. The method is useful for identifying an analyte
CC that modulates expression of a reporter sequence and in isolating and
CC identifying mouse and human transcriptional control elements associated
CC with cytochrome expression e.g. Cyp3A11 and Cyp3A4 respectively. The
CC cells and transgenic animals containing the expression cassette may be
CC used as models for studies involving toxicity and drug metabolism. The
CC present sequence is mouse Cyp3A11 promoter
XX Sequence 12275 BP; 3054 A; 2555 C; 2636 G; 4027 T; 0 U; 3 Other;

Query Match 21.3%; Score 486.098; DB 1; Length 12275;
Best Local Similarity 61.9%; Pred. No. 5.6;
Matches 1022; Conservative 0; Mismatches 509; Indels 121; Gaps 17;

Qy 416 GTCAGATAGGACATAGAGTATTTCATTTCTTTTATCTGTGAGACTTGTCTTGT 475
Db 2503 GTCGTATAGGATGATGGGACAAATTCATATTTTGTATCTGTTGAGCCTGTTTGTG 2562
Qy 476 TTGAATATGATTCATTTTGGAG-CTTTTCATAGGCTGCTGACAGAGGTACAGTC 533
Db 2563 ACAATTATATGTTAAATTTTGGAGAAGGTTCCGTGAGGTGCTGAGAATATAT---CAT 2619
Qy 534 TTTGTGTTTGGTGAATAGTCTGTAATATCTCT-AGGTCCACTTGGTTTATGACATCA 592
Db 2620 TTTGTTTAGGATAAAATGTTCTGTAGATATCTGCAATCCATTTGTTTCACTTCT 2679
Qy 593 GTTAGCTCAGCATTTCTCTGTTTCTGTTTGTGTTGAGATGACCTTACCTGTTGGAGAA 652
Db 2680 GTTAGTTTCA--CTGTGCTCTGTTTGTGTTTCTGTTTCAATGATCTGTCACATGATGA 2737
Qy 653 TGGGTTATGAAGTAGCCACCTATCTGCTGTGAGGT-CAATATGATGATTTAGCTGTAG 711
Db 2738 TGGTGTGTTGAAGTCTCCCACTATTTATTTGTGAGGTGCAATGTGCTTTGAGCTTTAC 2797
Qy 712 CTGTGCTTTGTTTATGAACCTTGGGTGACATTTGTTTGGTGCATAGACATTAAGAAATGC 771
Db 2798 TAAAGTGTCTTTAAATGAATGGCTGCCCTTGCANTTGGAGCATAGATATTCAAAATTGA 2857
Qy 772 AATGCTCTTGTGAGATTTT-CCTTTGTAGCTGATGATGATTTCTTCCCATCTCATCT 830
Db 2858 GAGTTCCTCTTGGAGATTTTACCTTTGATGATGATGAAGTGTCCCTCTCTTTT 2917
Qy 831 GCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAAGATATTAATGACCTGTATCGGCTTGC 889
Db 2918 GATAACTTTGTTTGAAGTGTGATTTTATTTGATATAGATGGCTTACCCAGCTTGT 2977
Qy 890 CTTAGGGCCATTTGCTTAGAATATC-TTTTCAATCTTTTACTCTAAGGTGATGCTATC 948
Db 2978 CTTAGACCAATTTGCTTGGAAAAATTTGTTTCCAGCCTTTCACTCTGAGGTAGTGTCTGC 3037
Qy 949 CAT---GGTAGGTTGCTTTTGGATGCAGAGTAGGATGGATCTTGTGTTTCATATCCA 1005
Db 3038 TTTTCCCTGAGATGGGTTTCTGTAGCAGCAGATGTTGGGCTCTGTTTGTAGCCA 3097
Qy 1006 TTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTTAGTATGTTTGA 1065
Db 3098 GTCTGTTAGTCTATGCTTTTATTGGGAATTTAG-----TCCATTGATTAAGA 3149

[illegible]

AC	AAF92531;
XX	
DT	16-MAY-2001 (first entry)
XX	
DE	Rat T2R04 nucleotide sequence SEQ ID NO:84.
XX	
KW	Human; rat; mouse; T2R; taste receptor; G-protein coupled receptor; taste transduction G-protein coupled receptor; identification; tongue; taste sensory neuron; taste cell; taste modulator; food; taste signalling pathway; ds.
KW	
XX	
OS	Rattus sp.
XX	
FN	W0200118050-A2.
XX	
PD	15-MAR-2001.
XX	
PF	08-SEP-2000; 2000WO-US024821.
XX	
PR	10-SEP-1999; 99US-00393634.
XX	
PR	22-FEB-2000; 2000US-00510332.
XX	
PA	(REGC) UNIV CALIFORNIA. (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Zuker CS, Adler JB, Ryba N, Mueller K, Hoon M;
XX	
DR	WFI; 2001-211396/21.
DR	P-PSDB; AAB87784.
XX	
PT	Nucleic acids encoding the T2R family of G-protein coupled taste receptors, useful for identifying taste modulators that can be used in food and pharmaceutical industries to customize taste, for e.g. to decrease the bitter taste of food.
PT	
XX	
FS	Claim 51; Page 195-197; 249pp; English.
XX	
CC	AAF92502 to AAF92572 represent nucleic acids which encode taste transduction G-protein coupled receptors designated T2R proteins. CC AAB87731 to AAB87824 represent T2R proteins, and AAB87825 to AAB87830 represent T2R family consensus sequences from the present invention. The T2R proteins are taste modulators. The nucleic acids are useful as probes for the identification of taste cells, as the nucleic acids are specifically expressed in taste cells. They also serve as tools for the generation of taste topographic maps that elucidate the relationship between the taste cells of the tongue and taste sensory neurons leading to taste centres in the brain. The taste modulators are useful for pharmacological and genetic modulation of taste signalling pathways. Modulatory compounds comprising T2R proteins can therefore be used in food and pharmaceutical industries to customise taste, for e.g. to decrease the bitter taste of food or drugs
XX	
SQ	Sequence 3449 BP; 795 A; 653 C; 702 G; 1299 T; 0 U; 0 Other;
	Query Match 20.9%; Score 475.598; DB 1; Length 3449;
	Best Local Similarity 63.0%; Pred. No. 11;
	Matches 965; Conservative 0; Mismatches 459; Indels 108; Gaps 16;
QY	589 ATCAAGTTAGTCGCAGCATTTCTGTGTTGGTTTTTGATGATGACCTAATGTTGGAG 648
Db	1115 ACCTGTTAGTTTCTGTGTCCTGTTTAG-TTTCGTTTCATGATCTGTCCAATTGATG 1173
QY	649 AGAATCGGCTATTGAAGTAGGCCACTPATCTGTGTGTGAGGTC-ANATGTGATTTTAGCT 707
Db	1174 AGCGTGGGGTGTGAAATCTCGACATATTGTTGTGTGAGATAAATGTGTGCTTTGAGCT 1233
QY	708 GTAGCTGTGCTGTTGTTTTATGAACCTGGGTGGACATTGTGTTGGTCATAGACATTAAAGA 767
Db	1234 TTAGTAAGATTTCCTTTTGTAATGTAGGTGCTTTTGCAITTCGTGCATAGATATTTRAGA 1293
QY	768 TTGCAATGTCCTCTTGGTGGGA-TTTTCCTTTGATGCCTATGTAGTATCTTCCCATACTC 826
Db	1294 TTGAGAGTTTCAGCTTGGTGGATTTTTCCTTTGATGAATGAAGTGTCTCTGTTATCTT 1353

QY 588 CATCAGTTAGTCCAGACATTTCTCTGTTTCGTTTTTTTGGTGGAGTACCTAACTGTTGGA 647
Db 14753 TTCTGTAAATTCACCTGTCTCTGTTTGTAGTTTG-TGTTTCCATGATCGCCCAATTGAT 14695
QY 648 GAGATGGGTATTGAAGTACCCACATCTGTGTGAGGT-CAATATGTGATTTTATGC 706
Db 14694 GAGAGTGGGTGTTGAAGATCCCACTAGTATGTGTAGGTGCAATGTGTGCTTGAGC 14635
QY 707 TGTAGCTGTCTGTTGTTTATGAACCTTGGTGACATGTGTGTGTGTCATAGACATTAAGA 766
Db 14634 TTTAATATGTTTCTTTTATTAATGTGGCTGCCCTTGCACTTGGATCATAGATGTTGAGA 14575
QY 767 ATTGCAATGTCTCTGTGTGATTTT-----CCTTTGATGCTATGTAGTATTTCTCCCA 821
Db 14574 ATTGAGAGTCTCTCTGTGTGATTTTATACCCCTTGATAAATATTCACGTCTTTCCCT 14515
QY 822 ATCTCATCTGCTTGTAGTTTGGGTTTAAGTCTAT-TAGTCAGATATTAAATGACTGTATC 880
Db 14514 AACITTTTGTAGTCTTTTGGTGAAGTCACTATATTCAATATTAGAATGGTTATTC 14455
QY 881 GCTTGTCTCTTAGGGCAATTTGCTTAGAATAT-CTTTTCCATCCTTTTACTCTAAGGTG 939
Db 14454 AGCTTGTCTCTTGGAAACCAATTCGCTTGGAAAATAGTTTTCAGACCTTTTACTCTGAGGA 14395
QY 940 ATGCTATCCATGGTAG--GTGTCTTTTGGATGACAGTAGGATGGATCTTGT 996
Db 14394 GTGTTGTCTTGACACTGAGATATGTTTCTGTATGACAGCAAAATGCTGGTCCATTT 14335
QY 997 TCATATCCATCTGTATCCCACTATCTTTTCTAGAGAAATAGATCATGAGTCATG 1056
Db 14334 ATGATCCAGCCTGTAGTCTATGTCTTTTATTGGGACTTGAGCCCAATTGA---TGTG 14278
QY 1057 ATGTTGAGAAATATCAATGACAGGTTTGTGGATCTTGTATCTTGCACTGTGAGT 1116
Db 14277 ATGTTTAAAGATATTAAAGATAGATGTTTGTCTCTGTTATTTTGTGTTAGTAT 14218
QY 1117 GT 1176
Db 14217 GGACTTATGTTTGTGTGTGTATATCTTTTGGTTGT----- 14180
QY 1177 GTCTGT 1236
Db 14179 -----TGAAGAAGATTAATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 14125
QY 1237 TTATTTATTTATTCATTTTCTTGAATGTGGTAACATCTTTAGATGAAGTTTTTCTCC 1296
Db 14124 TTTTACATCTATTAT----- 14110
QY 1297 TAGCCTTCTTTAGTCTGCAATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGA 1356
Db 14109 -----CCTCTGTAGGACTGCAATTTGTTTAAATCTATTTGTTTAAATTTGGTTTGTGATGGA 14054
QY 1357 ATGCTCTTCTTCTCCAACTATTTGACAGAAAGTTTCTTAAAGTGCAAGTGTCTGGCT 1416
Db 14053 ATATCTGTTTCTCCATCTATGGAATTTGAGATTTTCTGGGTATAGTACCTAGCT 13994
QY 1417 GACATCTGTAGTCTCTTGAGTCTGTAGCACATCTGTGCGAGGCTTCTTACATTTTGA 1476
Db 13993 GGCAATTTATGTTCTCTTAGGCTCTGTATGAGATCTGCCAGTATCTTCTAACTTTTCATAG 13934
QY 1477 TTTCTATTGGAAGTCTAGGTGTTATCTTATACATCTGCTTATATATGTTAAATTTGTTCT 1536
Db 13933 TCTCTGGT-GAAGAGTCTGGGTAAATATGATAGTCTGCTTAAATTTTAAATTTTATTTGACCT 13875
QY 1537 TTTTTCCTTGCATCTTTTAAATCTTTTCTTGTCTTATCTTATCTTTTACTGATTTTGAATAT 1596
Db 13874 TTTTCCCTTACTACTTTTAAATCTTTCTTTGTTTGTGCAATTTAGTATTTTGGATAT 13815
QY 1597 TATGACTGTGGGAGTTTCTTTCCGTCCTCAATCTA-TTTGGTGTGTTTGTATGCTTCT 1655
Db 13814 TATGTGACAAAGGAATTTCAATTTCTGCTTCTATCTACTTTGGAAATCTGTAGGCTTCTT 13755

QY 1656 GTACCTTGATAGGACTCTCTTCTCAAGGTAGGAAATTTTCTTTTGGTTTCTTTCTTGA 1715
Db 13754 GT-----ATGGGCACTCTTCTTTTAGTTAGGAAGTTTCTTCTATATTTTATGA 13701
QY 1716 AAATATTTTCCCTGCTTTTGTGACCTGCCCTTCTTCCCTTCCCT----- 1756
Db 13700 AGATATTTACTGGCCCTTTAACTTGGGAATCTTCACTCTCTTCTTTAGGTATAGAAGAAC 13641
QY 1757 -----CTATTCCCTTGGTTTTCATAGTGTCTCTGGC 1789
Db 13640 TATCCTTATAGGCTGTGTTCTTATATTTATCTTATCTTAGGCTTCTTATGTGCTGTGAT 13581
QY 1790 TTCTCGGATGTTTATGCTGCTGATTTATTTAGACTTAAATTTCTTTGACCAAGGTATC 1849
Db 13580 TTCTCGGATGTTTGGGATAGGAATTTTGTATTTGCAATTTCTTTGACTGTGTGTC 13521
QY 1850 CATTTCTTCTATCTTGTCTTCACTGCTGAGATTTCTCTCTCTATCTCTTGTATTTCTGTC 1909
Db 13520 AATGTTTCTATGGTATCTTCTGACCTGGGATCTCTCTCTATCTCTTGTATTTCTGTT 13461
QY 1910 AGTGAGGCTGTCTCTGAGGTTCCCTG-ITGGGTTCTTAAATTTTCAITTCAGATTTCC 1968
Db 13460 GGTGATGCTTATCATATGACTCTCTGATCTCTTCTAGGTTTCTATCTCCAGGTTGG 13401
QY 1969 TTCAGTTTGGGTTTGTGTTTATTTAAATTTTCCACTTTCAGGCTCTGAAATGTTTACT 2028
Db 13400 CTCCCTTGTGATTTCTTTTATTTGTTCTCTATTTCCATCTTATGATCTCTCAATGGTTTGT 13341
QY 2029 CATTTTCTCT---CCCAGTATTTACATTTTTCATAGGTTTCTTTAAATGGATTTAT 2078
Db 13340 CAGTCTCTTACCTGTTTGGTTTGTGTTTCTGTAATCTTTAAGGATTTT 13288

RESULT 18

ADA02777/c

ID ADA02777 standard; DNA; 96595 BP.

XX ADA02777;

XX 06-NOV-2003 (first entry)

XX Mouse Ptpkr carcinoma associated gene, SEQ ID NO:1295.

XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
XX gene; ds.

XX Mus sp.

XX WO2003057146-A2.

XX 17-JUL-2003.

XX 26-DEC-2002; 2002WO-US041414.

XX 26-DEC-2001; 2001US-00035832.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-587068/55.

XX New recombinant nucleic acid encoding carcinoma associated protein,
XX useful for preparing compositions for treating carcinomas.

XX Claim 1; SEQ ID NO 1295; 245pp; English.

XX The invention relates to recombinant carcinoma associated (CA) nucleic
XX acid sequences from mouse and human (ADA01482-ADA03094), and to
XX recombinant carcinoma associated proteins (CAP) encoded by them. The
XX invention also encompasses expression vectors and host cells comprising a
XX CA nucleic acid, a polypeptide (especially an antibody) that specifically

CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed murine CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 96595 BP; 27994 A; 18068 C; 18732 G; 30838 T; 0 U; 963 Other;
Query Match 19.9%; Score 453.098; DB 1; Length 96595;
Best Local Similarity 60.2%; Pred. No. 3.1;
Matches 1007; Conservative 0; Mismatches 509; Indels 157; Gaps 17;
Qy 470 TTTCTTTTGAATAATGATTCATTTTGGAGA--GTTTCTATAGGGTCTGCACAGAAGGT 527
Db 14867 TTAATAATTTAGGTATAGTCAGTTTGGAGAATACCAATAGGTCTCAGAAGATGTT 14808
Qy 528 ACAGCTTTGGTTTTCGGTGAATAGTCGTAATATCTCTAGTCCACATGGTTTATGA 587
Db 14807 ATATCTTTTATTTAAGATGAATAGTTATATATATATAT-ATTTGGTCCATAA 14754
Qy 588 CATCAGTTAGTCCAGCATTTCTCTGTTTGGTTTGTGAGATGACCTAATCTTTGGA 647
Db 14753 TTTCTGAAGTTTCACTGTGTCTCTGTTTGTAGTTG-TGTTTCCATGATCTGCCATGAT 14695
Qy 648 GAGATGGGGTATGAAGTAGCCCACTATCTGTGTGTGAGGT-CAATATGATTTATGC 706
Db 14694 GAGATGGGGTATGAAGATCCCACTAGTATTTGTTAGGGTGAATGTGTGCTTGAGC 14635
Qy 707 TGTAGCTGTGCTGTTTATGAATCTGGGTGACATTTGTTTGGTGATAGACATTAAGA 766
Db 14634 TTTAATATGTTTCTTTTATTAATGTTGGCTGCCCTTGACATCATAGATGTTTCAGA 14575
Qy 767 ATTGAATGTCCTCTTGGTGAATTT-----CCTTTGAGGCTATGATGATTTCTTCCCA 821
Db 14574 ATTGAGAGTCTTCTTGTGTAGATGTTTACCTTTTGATTAATATCAAGTCTTTCTCT 14515
Qy 822 ATCTCATCTGCTAGTTTGGTGTATAGTCTPAT-TAGTCAGATATTAATAATGACTGTATC 880
Db 14514 AACTTTTTGTAGACTTTTGTGTAAGTACTATATATCAATATAGATGGTTATTC 14455
Qy 881 GGCTTGCTTTAGGGCCATTTGCTTAGAATAT-CTTTTCCATCCTTTTACTCTAAGGTG 939
Db 14454 AGCTTGTCTTGGAAACCATTCGTTGGAAATAGTTTCCAGACCTTACTCTCGAGTA 14395
Qy 940 ATGCTATCAATGGTAG---GTTGCTTTTGTGATGACGACAGTAGGATGATCTGTTT 996
Db 14394 GTGTTGTCTTGGACACTGAGATAGTTTCTGTATGACGAAATGCTGGGTCCATTT 14335
Qy 997 TCATATCCATTTCTGTATCCAGTATCTTTTCTAGAGAAATTAAGATCAITGACTCATG 1056
Db 14334 ATGTATCCAGCTGTAGTCTATGCTTTTATTTGGGACTTGAGCCCATGA---TG TG 14278
Qy 1057 ATGTTGAGAAATTAATCAATGACAGGTTTGTGGATTCCTTGTATCTTGCACATTTGAGT 1116
Db 14277 ATGTTAGAGATATTAAGGAATAGTATGTTGTCTTCTGTTATTTTGTGTTAGTGAT 14218
Qy 1117 GT 1176
Db 14217 GAACCTATGTTGT 14180
Qy 1177 GTCTGT 1236

RESULT 19

ADB72515/c

ID ADB72515 standard; DNA; 96595 BP.

XX

ADB72515;

XX AC

XX 04-DEC-2003 (first entry)

XX DE

Mouse Pcpk gene.

XX mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
 KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
 KW Mus sp.
 OS WO2003008583-A2.
 FN 30-JAN-2003.
 XX 26-DEC-2001; 2001WO-US051291.
 PF 02-MAR-2001; 2001US-00798596.
 PR 23-OCT-2001; 2001US-00004113.
 PR 08-NOV-2001; 2001US-00052482.
 PR 30-NOV-2001; 2001US-00997722.
 PR 20-DEC-2001; 2001US-00034650.
 XX (SAGR-) SAGRES DISCOVERY.
 PA Morris DW, Engelhard EK;
 XX WPI; 2003-239337/23.
 XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
 PT cancers, neoplasm, adenocarcinoma, or sarcomas.
 XX Claim 1; SEQ ID NO 343; 2304pp; English.
 CC The invention relates to a novel recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the 660 sequences fully defined
 CC in the specification. A polynucleotide of the invention has cytostatic
 CC activity, and may have a use in gene therapy, or in a vaccine. The
 CC recombinant nucleic acids and polypeptides are useful for treating
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
 CC sarcomas. The present sequence represents a mouse gene of the invention.
 XX Sequence 96595 BP; 27994 A; 18068 C; 18732 G; 30838 T; 0 U; 963 Other;
 SQ
 Query Match 19.9%; Score 453.098; DB 1; Length 96595;
 Best Local Similarity 60.2%; Pred. No. 3.1;
 Matches 1007; Conservative 0; Mismatches 509; Indels 157; Gaps 17;
 470 TTTCTTTGAATATGATTTCAATTTGGAGA--GTTTCATAGGCTCTCACAGAGGT 527
 14867 TTACTATATTAGTATAGTCAGTTTGGAGAGATACCATAAGGTCTCAGAGATGGT 14808
 528 ACAGTCTTTGTTTGGTGAATAGTCTGTAAATATCTCTAGTCCACCTTGGTTATGA 587
 14807 ATATCTTTTATTAAGGATGAATGTTATATAATATAT--ATTTGTCATTA 14754
 588 CATCAGTAGTCCAGCATTTCTCTGTTTGGTTTGGATGACCTAACTGTGGA 647
 14753 TTTCTGTAAGTTTCCACGTGCTCTGTTAGTTG-TGTTTCCATGATCTGCCCATTTGAT 14695
 648 GAGAAATGGGTTAAGTAGTACCCACATCTGTGTGTGAGGT-CAATATGTGATTTAGC 706
 14694 GAGAGTGGGTTTGAAGATCCACATAGTATGTGTAGGTTGCAATGTGCTTGAGC 14635
 707 TGTAGCTGTGCTGTTTATGAATCTGGTGACATTTGTTGGTGATAGACATTAAAG 766
 14634 TTTAATATGTTTCTTTTATTAATGTGGCTGCCCTTGCACCTTGGATCATAGATTTTCA 14575
 767 ATTCGAATGCTCTCTGGTGATTTT-----CCTTTGATGCCATGTAGTATTTCCCA 821
 14574 ATTGAGAGTCTCTTGGTAGATTTTACCCTTTGATATAAATCAACGCTTTCTCTT 14515
 822 ATCTCATCTGCTTAGTTTGGGTTAAGTCTAT-TAGTCAGATATTAAATGACTGTATC 880
 14514 AACTTTTTTGAAGACTTTTGGTTGAAAGTCACTATTATTCAATATTAGAATGTTATTC 14455
 881 GGCTTGCTTCTTAGGGCCATTTGCTTAGAATAT-CTTTTCCATCCCTTTACTCTAGGTG 939

14454 AGCTTCTTTCTTGGAACCACTTCGCTTGGAAATAGTTTTCCAGACCTTTTACTCTCAGGTA 14395
 940 ATGTCTATCCATGGTAG---GTTTCTCTTTTGGATGAGAGTAGGATGATCTGTGTT 996
 14394 GTGTTTGTCTTTTGACACTGAGATATGTTTCCCTGTATGACAGAAATGCTGGGTCCATTTT 14335
 997 TCATATCCCACTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATTG 1056
 14334 ATGTATCCAGCCCTGTAGTCTATGTCATGCTTTTATTCGGGACCTGAGCCCATTTGA 14278
 1057 ATGTTGAGAAATATCAATGAGCAGTGTGTTGGGATCTTGTATCTTGTGACTTTGCAAGT 1116
 14277 ATGTTAAGAGATATTAAAGGAATAGTGTGTTGCTTCTGTTATTTTGTGTTAGTGTAT 14218
 1117 GT 1176
 14217 GGACTTATGTTGTGTGGCTATATCTTTTGGGTTTGT----- 14180
 1177 GTCTGTGTGTGTCTGT 1236
 14179 -----TGAAGAAGATTAATCTTCTGCTTTGCTAGGGGTAGTGTGCTTGTGTTGGAA 14125
 1237 TTATTTATTAATCATATTTCTTGTGAATGTTGGGTACATCTTTAGATTTGAAGTTTCTCC 1296
 14124 TTTTACATCTATTAT----- 14110
 1297 TAGCCTTCTTTAGGTCCTGCAATTTGAAGATAGATATTTCTTTACATCTGATTTTATCTAGA 1356
 14109 ---CCTCTGAGGACTGCAATTTGTTAAACTATGTTTAAATTTGTTTGCATGGA 14054
 1357 ATGCTTCTTCTTCCAACTATGTGACAGAAAGTTTCTTAAGTCAGTAGTCTGGCCT 1416
 14053 ATATCTTGTGTTTCTCCATCTATGTTAATGTAGAGTTTGTGGGTATAGTAGCCTAGGCT 13994
 1417 GACATCTGTAGTCTCTGAGTCTGTAGCATCTGTGACGGGCTCTTACATTTTGTAG 1476
 13993 GGCATTTATGTTCTTTAGGGTCTGTATGAGATCTGCCAGATCTTCTACTTCTCATAG 13934
 1477 TTTCTATTGGAAAGTCAGGTGAATTTCTATATACATCTGCTTTTATATGTTAAATGTTGCT 1536
 13933 TCTCTGGT-GAGAAGTCTGGGTAATATGATAGTCTGCTTCTTAAATGTTTATTTGACCT 13875
 1537 TTTTTCCTCTGCATCTTTTAAATATCTTTCTTCTTGTCTATACTTCTTACTGATTTGATTTAT 1596
 13874 TTTCCCTTACTACTTTTAAATTTCTTTCTTGTGTCATTTAGTATTTAGTATTTAGTAT 13815
 1597 TATGACATCTGGGGAGTTTCTTTCCGGTCCAATCTA-TTTGGTGTGTTTGTATGCTTCTT 1655
 13814 TATGTGACAAGAGGAATTTCAATTTCTGCTTCTATCTACTTTTGAATCTGTAGGCTTCTT 13755
 1656 GTACCTTATAGGATCTCTTCTCAGAGTTAGGAATTTTCTTCTTTTGGTGTCTTGA 1715
 13754 GT-----ATGGGCACTCTTTTCTTAGGTTAGGAAGTTTCTTCTATAAATTTTGA 13701
 1716 AAATATTTTCCCTGTCTTTTGAACCTGCTTCTTCCCTTCCT----- 1756
 13700 AGATATTACTGGCCCTTTTAACTTTGGGAATCTTCACTCTCTTTTAGTATAGAAGAAC 13641
 1757 -----CTATTCCTTGGTGTGTCATAGTGTCTCTGGC 1789
 13640 TATCCTTATGCTTGTCTTCTTATTTATCTTATCTTATGCTTCTTATGCTCTCGAT 13581
 1790 TTTCTGGGATGTTTATGCTGCTGATTTTATAGACTTAACATTTTCTTGTGACCAAGGTATC 1849
 13580 TTCTGGATGTTTGGGATAGGAATTTTGTATTTTGGCATTTTCTTGTACTGTGTGTC 13521
 1850 CATTTCTTCTATCTGTCTTCACTGCTGAGATCTCTCTTCTATCTCTTGTATCTGTGTC 1909
 13520 AATGTTTCTATGATCTTCTGACCTGGGATCTCTCTTCTATCTCTTGTATCTTGT 13461
 1910 AGTAGGCTTCTCTGAGGTTCTGTG-TTGGGTTCTTAAATTTTCAATTTTCCAGATTTCC 1968
 13460 GGTGATGCTTACATCTATGACTCTCTGATCTCTTCTTCTAGGTTTCTATCTCCAAGTTGG 13401

QY 1361 CTTCTTTCTCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCTGACA 1420
 DB |||||
 QY 39325 CTTGGTTTCTCATCTATGGTAATGATGTTTCTGGGTATAGTAACCTGAGGTGGCA 39266
 DB |||||
 QY 1421 TCTGTAGTCTCTTGGAGTCTCTAGCACATCTGTGCAGGGCCTCTTACATTTGAGTTTC 1480
 DB |||||
 QY 39265 TTTGTGCTCTCTTATGGTCTATATGACATCTGCCAGGATCTTCTAGCTTTCATAGTCTC 39206
 DB |||||
 QY 1481 TATTGGAAGTCAAGTCTAATCTTATACATCTGCCCTTATATGTTAAATGGTCTTTT 1540
 DB |||||
 QY 39205 T-GGGGGAAGTCAAGTCTAGTTTGTATAGTCTGGCTTTATTTGTTACTTGAAC-TTCT 39148
 DB |||||
 QY 1541 TCCCTTGCATCTTTTAATATCTTTCTTTGTTCTATACCTTTTAGTGATTTGATTTATG 1600
 DB |||||
 QY 39147 TCCCTTACTGCTGTAGT- ---CGTCTTTGTTGTGCAITTTGTTGATAATATG 39092
 DB |||||
 QY 1601 CACTGTGGGGAGTTCTTTTCCGGTCCAACTATTTGGTGTTTGTATGCTTCTTGTAAC 1660
 DB |||||
 QY 39091 TGACAGGAGGAATTTCTTTTCTGGTCCAACTATTTAGAGTTCTGTAGGCTTCTTGTATG 39032
 DB |||||
 QY 1661 TTGATAGGCATCTCTTCTCAAGGTTAGCAAAATTTCTTTTGGTTTCTTGAAATA 1720
 DB |||||
 QY 39031 TTCAATGGAATCTTTCTTTAGTTAGGAGAGTTTCTTCTATAATTTTGTGGAAGATA 38972
 DB |||||
 QY 1721 TTTTCCCTGCTTTTGACCT- GCCTTCTTCCCTTCCCTCTATCTTCTTTGGTTTTCATAG 1779
 DB |||||
 QY 38971 TTTTACTGGCCCTTTAAATGGCAATCTTCTTCTGTCTATACCTATTTCTTTAGGTTT 38912
 DB |||||
 QY 1780 TGCTCT 1784
 DB |||||
 QY 38911 GGTCT 38907
 DB |||||

RESULT 21

ID ADB72575/c
 XX ADB72575 standard; DNA; 41637 BP.
 AC ADB72575;
 XX
 DT 04-DEC-2003 (first entry)
 DE Mouse Map3k8 gene.
 XX
 DE mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
 KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
 XX
 OS Mus sp.
 XX
 PN WO2003008583-A2.
 XX
 PD 30-JAN-2003.
 XX
 XX 26-DEC-2001; 2001WO-US051291.
 XX
 XX 02-MAR-2001; 2001US-00798586.
 PR 23-OCT-2001; 2001US-00004113.
 PR 08-NOV-2001; 2001US-00052482.
 PR 30-NOV-2001; 2001US-00937722.
 PR 20-DEC-2001; 2001US-00034650.
 XX
 XX (SAGR-) SAGRES DISCOVERY.
 XX
 XX Morris DW, Engelhard EK;
 XX WPI; 2003-239337/23.
 DR
 XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
 PT cancers, neoplasm, adenocarcinoma, or sarcomas.
 XX
 XX Claim 1; SEQ ID NO 403; 2304pp; English.
 PS
 XX The invention relates to a novel recombinant nucleic acid comprising a
 CC nucleotide sequence selected from any of the 660 sequences fully defined

CC in the specification. A polynucleotide of the invention has cytostatic
 CC activity, and may have a use in gene therapy, or in a vaccine. The
 CC recombinant nucleic acids and polypeptides are useful for treating
 CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
 CC sarcomas. The present sequence represents a mouse gene of the invention.

XX
 SQ Sequence 41637 BP; 12049 A; 8888 C; 9032 G; 11648 T; 0 U; 20 Other;

Query Match 16.8%; Score 381.798; DB 1; Length 41637;
 Best Local Similarity 60.7%; Pred. No. 8.7;
 Matches 914; Conservative 0; Mismatches 467; Indels 124; Gaps 19;

QY 292 TCTATTCTTGATTTCTATCTCTGGCTCATTTTAACTCAGTAGTGAG-TTGTTCGTTTC 350
 DB |||||
 QY 351 CATAGTTTGTAAAGTTTCTGTTGTTCTGTTGTTGTTGTTTATCTAGATTAAAGCTG 410
 DB |||||
 QY 40239 CACATGATGTGAGCTTCTATAATATGTTGTTAGAGATCAGCC---TTAGTCT 40184
 DB |||||
 QY 411 TGGTGTCTAGATAGGACATAGAGTATATTTCAATGTCTTTTATCTGTGAGACTTGTCT 470
 DB |||||
 QY 40183 TGGTGATATGATAGAAATGCAATGGGATTTTCAA--TCTTATATCTGTGAGGCTGTT 40127
 DB |||||
 QY 471 TTGTTTGAATATGATATCAATTTTGGAGA--GTTTCATAGGTGCTCACAAGAGATA 528
 DB |||||
 QY 40126 TTGTGACCTATTATATGGTCACTTTTGCAAAAGGTACCTGAGGTGCTGAGAGAGATA 40067
 DB |||||
 QY 529 CAGTC-TTTGTGTTTGTGGTGAATATAGTCTGTAAATATCT-CTAGTCCACTTGGTTATG 586
 DB |||||
 QY 40066 TACCCCTTTGTTTAGGATAAAATGTTCTATAGATATCTGTTAAATCCATTGTTTCA 40007
 DB |||||
 QY 587 ACATCAGTAGTCTCAGCAATTTCTCTGTTTCGTTTGTGTTGTTGTTGTTGTTGTTGTTG 646
 DB |||||
 QY 40006 ACTTCTGTTAGTTTTCACCTGTCTGTCTGTCTGTTTCTCATT-CTATGATCTGTCCATTGA 39948
 DB |||||
 QY 647 AGAGATGGGTATTGAAGTAGCCACTATCTGTGTGTG-AGGTCAATATGTGATTTTAG 705
 DB |||||
 QY 39947 TGAGAGTGGGCTATTGAAGTCTACCACTATTTATGTTGTCAGTCTGTGTTGTTGAG 39888
 DB |||||
 QY 706 CTGTAGCTGTCTGTTTATGAACCTGGGTGACATTTGTTGTTGTTGTTGTTGTTGTTGTTG 765
 DB |||||
 QY 39887 CTTTAGTAAACTTCTTAAATGAAATGTGGAATTCCTTGCATTGAGATATAGATCTCAG 39828
 DB |||||
 QY 766 AATTGCAATGCTCTCTGTTGGATTTTCTTTGATGCTCTATGATTTCTTCCCAATCT 825
 DB |||||
 QY 39827 ATATTGAGTTCATCTTGTGATGTTTATCTT--TGATGATGAAGTTTCCCTTCTTCT 39770
 DB |||||
 QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA--TTAGTCAGATATTAATAAGTCTGATCGGC 883
 DB |||||
 QY 39769 TTTTGTGATAAATTTTAGGTTGAAATTTCAATTTTATTTATAGAAATGGTACTCCAG 39710
 DB |||||
 QY 884 TTGCTTCTTAGGGCCATTTGCTTAGNATA-TCTTTTCCATCTCTTACTCTAAGTGTGT 942
 DB |||||
 QY 39709 TTGTTTCTTGGGACCAATTTGCTTGGAAAATGTTTTCATCTTCTGCTCGGGTAGCT 39650
 DB |||||
 QY 943 TCTATCCATGTTAGTGTGCTTTTGTGATGACAGTAGGATGATCTGTTTTCATAT 1002
 DB |||||
 QY 39649 TCTGT-CTTTGTCACTGAGGTGGTGTATGACAGCAAAATTTGGTCTCTGTTATGTAA 39591
 DB |||||
 QY 1003 CCATTCGTACCAGTATCTTTTCTAGAGAAAATTAAGATCATTTGAGTCATGATGTTG 1062
 DB |||||
 QY 39590 CCAGTCTTTTAGTCTATGATTTATTTAGG-----TCCATTGATATTA 39547
 DB |||||
 QY 1063 AGAATTATCAATGAGCAGTGTGTTGTTGATCTTGTATCTTGCACTTGTGAAGTGTGT 1122
 DB |||||
 QY 39546 AGAGATATTAAGGAAGATCATTTGTTCTCTGTTATTTTGTGTTATAGTTGAATT 39487
 DB |||||
 QY 1123 GT 1182
 DB |||||
 QY 39486 CTGTTTCATGTG----- 39476
 DB |||||
 QY 1183 GTCTGTGTCTGTGTGTGTGTGTGTCTCTCCCTCTTTTTCATTTTGGCCTGGAATATT 1242
 DB |||||

```
Db 39475 -----CCTATCTCTTTTACGTCGTGAAGATTA 39446
Qy 1243 ATTATTCATATTTCTTGAATGCGGTAAACATCTTTAGANTGAAGTTTCTCCTAGCCT 1302
Db 39445 CTTTCTGCTTGTCTAGGCATAGATTCCTCTTGTGTTGAAGTTTCCCTTTATTAT 39386
Qy 1303 TCTTTT--AGGTCCTCATTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAATGT 1360
Db 39395 CTTTGAAGGCTGGAATTTGGAAGATATTTTGTAAATTTGTTTGTTCATGGAATAG 39326
Qy 1361 CTTTCTTCTCCAACTATTGTGACAGAAAGTTTCTTAAAGTGCAGTAGTCTGGCCTGACA 1420
Db 39325 CTTGCTTCTCCATCTATGTAATGATAGTTTGTCTGGGTATAGTAACCTGAGGTGGCA 39266
Qy 1421 TCTGTAGTCTCTTGGAGTCTCTAGACATCTGTGCGAGGCGCTTCTTACATTTGAGTTTC 1480
Db 39265 TTTGTGCTCTTATGGTCTATATGACATCTGCCCGAGATCTTCTAGCTTTTCATAGTCTC 39206
Qy 1481 TATTGGAAGTCAAGTGTAAATCTAATACATCTGCCCTTTATATGTTAAATGGTCTTTT 1540
Db 39205 T-GGGGAAAGTCAAGTGTAGTTTGTAGTCTGCCCTTTATTTGTTACTTGAAC-TTCT 39148
Qy 1541 TCCCTTGCACTTTTAAATATCTTCTTGTCTTATCTTCTTGTCTATCTTTTACTGATTTATATG 1600
Db 39147 TCCCTTACTGCTTAGT---CGTCTTGTGTTGTCATTTGGTGTGTTTGATAAATATG 39092
Qy 1601 CACTGTGGGAGTTTCTTTCCGGTCCCAACTATTTGGTGTGTTTGTATGCTTCTTGTACC 1660
Db 39091 TGACAGGAGGAATTTCTTTTCTGTGCCAACTATTTAGAGTTCGTAGGCTCTTGTATG 39032
Qy 1661 TTGATAGCATCTCTTCTCAAGGTTAGGAAATTTTCTTTTGGTCTTCTTGAAATA 1720
Db 39031 TTCAATGTAATCTCTTCTTTAGTTAGGTAAGTTTCTTCTATAAATTTGTTGAAGATA 38972
Qy 1721 TTTTCCCTGCTTTTGACCT-CCCTTCTCCCTTCTCTATCTCTTGTGTTTGTGATAG 1779
Db 38971 TTTACTGCCCTTTAAATGGAATCTCTCTTCTGTCTATACCATATATATCTTAGT 38912
Qy 1780 TGTCT 1784
Db 38911 GGTCT 38907

RESULT 22
ADC85316/c
ID ADC85316 standard; DNA; 41637 BP.
AC ADC85316;
XX
XX 01-JAN-2004 (first entry)
XX Human Mef2c coding sequence.
XX
XX Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
XX secreted; transmembrane; intracellular; ds.
XX Homo sapiens.
XX
XX WO2003045230-A2.
XX
XX 05-JUN-2003.
XX
XX 02-DEC-2002; 2002WO-US038582.
XX
XX 30-NOV-2001; 2001US-00997722.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX PI Morris DW, Engelhard EK;
XX WPI; 2003-513603/48.
XX
```

```
PT New recombinant nucleic acid comprising a nucleotide sequence of any of
PT the carcinoma-associated (CA) genes, useful for screening for drug
XX candidates for diagnosing or treating carcinomas.
XX
XX Claim 1; SEQ ID NO 102; 983pp; English.
XX
XX The invention relates to a recombinant nucleic acid comprising a
XX nucleotide sequence selected from any of the fully defined carcinoma-
XX associated (CA) genes from the 50 tables given in the specification. The
XX CA proteins are secreted, transmembrane or intracellular proteins. The
XX recombinant nucleic acids are useful for screening for drug candidates
XX for diagnosing or treating carcinomas. Sequences given in ADC85215-
XX ADC85514 represent CA genes of the invention.
XX
XX Sequence 41637 BP; 12049 A; 8888 C; 9032 G; 11648 T; 0 U; 20 Other;
XX
XX Query Match 16.8%; Score 381.798; DB 1; Length 41637;
XX Best Local Similarity 60.7%; Pred. No. 8.7;
XX Matches 914; Conservative 0; Mismatches 467; Indels 124; Gaps 19;
XX
Qy 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAG-TTGTTCGTTTC 350
Db 40299 TTTTTCCTTTATTTTTCCTTGACCAATTTATCATGTAGTAGAGTGTGTTCACCTTC 40240
Qy 351 CATAGCTTTGTAAAGTTTCTGTTTCTGTTGTTGTTGTTATCTAGATTAAAGCTG 410
Db 40239 CACATGATGTGAGCTTTCTATTAATATGTTGTTAGAGATCAGCC---TTAGTCCT 40184
Qy 411 TGTGTCTAGATAGGACATAGAGATATTATTTCAATGTCTTTTATCTGCGAGACTTGCT 470
Db 40183 TGTGTATATAGATAGATGATGCGGATTAATTTCAA---TCTTATATCTGTGAGGCTGT 40127
Qy 471 TTGTTTGAATATGATPATTCATTTTGGAGA--GTTTCATAGGTCCTGCACAGAGGTA 528
Db 40126 TTGTGACCTATTATATGTCAGTTTTCGAAAAGTACCTGAGTGCTCGAAGAGGTTA 40067
Qy 529 CAGTC-TTGTGTGTTTGGTGAATAGTCTGTAATAATCT-CTAGGTCCACTTGGTTATG 586
Db 40066 TACCCCTTTGTTTAGGATAAAATGTTCTATAGATATCTGTTAAATCCATTTGTTTCATA 40007
Qy 587 ACATCAGTTAGTCCAGCAATTTCTCTGTTTCGTTTCTGAGATGACCTAACTGTTCG 646
Db 40006 ACTTCTGTTAGTTTCACTGTGTCCTGTTCAAGTTCTCAIT-CTATGATCTGTCCATGA 39948
Qy 647 AGAGATGGGTATTGAAGTAGCCCACTATCTGTGTGTG-AGGTCAATATGTGATTTTAG 705
Db 39947 TGAGAGTGGGCTATTGAAGTCTACCACTATTATTGTGTCAGTGTCTATGTTTGTAG 39888
Qy 706 CTGTAGCTGTGCTGTTTATGAACTTGGGTGACATTTGTTGGTGCAATAGACATTAAG 765
Db 39887 CTTTAGTAAACTTTCTTTAATGAATGTGAATTCCTTGCATTTGGAGTATAGATGTTTCA 39828
Qy 766 AATTGCAATGTCTCTTGTGTTGATTTTCCCTTGTAGTCCCTATCTAGTATTCTTCCCAATCT 825
Db 39827 ATATTGAGTTTCACTTGTGATATTTATCTT--TGAGTATGAAAGTTTCCCTTCTTATCT 39770
Qy 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA--TTAGTCAGATATTAAAGTCTGTTATCGGC 883
Db 39769 TTTTGTGATAACTTTAGGTTGAATTTCAATTTTATTATTATTAGATGGCTACTCCAGC 39710
Qy 884 TTGCTTCTTAGGGCCATTTTGTCTTAGAATA-TCTTTTCCATCCTTTTACTCTAAGGTAG 942
Db 39709 TTTTTCITGGGACCATTTGCTTGGAAATTTGTTTCCATCTTTTGTCTCTGGGTAGCT 39650
Qy 943 TCTATCCATGGTAGGTTGTCTTTTGTGATGAGCAGTAGGATGGATCTTGTGTTTTCATAT 1002
Db 39649 TCTGT-CTTTGTCACGTGAGGTGGTTGTATGACAGAAATTTTGGTCTCTGTTTATGTA 39591
Qy 1003 CCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCATGTTGTTG 1062
Db 39590 CCAGTCTTTTACTCTATGTAATTAATTAG-----TCCATTGATATTA 39547
Qy 1063 AGAATTATCAATGAGCAGGTTTGTGGATTCTTGTATTCTTGCACCTTGTGGAAGTGTGT 1122
```



```

Db 1425 AGTTTTCTCTATAAATTTGTTGAAGAATATTGTCGCCCTTTAAAGTTGAGAACTTCA 1366
Qy 1741 ----CCITCTCCCTCCTCTATTCCCTTTGGTTTTGTCATAGTGTCTCTGGCTTCCTGG 1796
Db 1365 TTCTCATCTACTCTATTATCCACAGGTTTGGTCTCTCTCAATTTGCTCGAATTTTGG 1306
Qy 1797 ATGTTTATGCTGGAATATTATAGACATTAACATTTCTTTGACCAAGGTAATCCATTTCT 1856
Db 1305 AGTTTTGAGTTAGGATCTTTTGGTTTTGCGATTTCTTTGGTTGTGTGCCGATGTC 1246
Qy 1857 TCTATCTTGTCTTCACTGCCCTGAGATCTCTCTCTATCTCTCTGTTATTTCTGTCAGTCAGG 1916
Db 1245 TCTAAGGAATCTTCTGACCTGAGATCTCTCTCTCACTCTCTGATTTCTGTCGATG 1186
Qy 1917 CTGTCCTGAGGTTCCG-TGGGTTCTTAAATTTTCAATTTCAATTCAGATTTCTCTCAGTT 1975
Db 1185 CTCACATCTATGTTCCAGATTTCTTTACTAGGTTTCTATCTCTAGGTTGTCTCAAT 1126
Qy 1976 TGGGTTTGTGTTTAAATCTTATTTCCACTTTCCAGTCCCTGAATGTTTACTCAATTTT 2035
Db 1125 TGGTTTTCTTGTG-TTCTACATCTCCCTTTTAGGTTCTGGATGATTTTGTTCATTTCC 1069
Qy 2036 CTCACGATTTACATTTTCAATTTTCTTTAATGGAATTTATTTCAATTTCTCTCAAG 2095
Db 1068 ATCACCCTGTTGGTTGCTTTTCTGTAATTTCTTTAAGAACTTCTACCTCTTTAGCAGTG 1009
Qy 2096 ACCTTTATGAAATCATAAATGATGATGTAAGGTCCTT 2133
Db 1008 TTCTCCTTACTCTTTTAAAGTGAGTTAATAAGCCCTT 971

```

RESULT 24

AAx83006

ID AAX83006 standard; DNA; 16442 BP.

XX AC AAX83006;

XX DT 31-AUG-1999 (first entry)

XX DE Partial mouse WRN genomic sequence #2.

XX KW Mouse; WRN; Werner's syndrome; detection; diagnosis; autosomal;

XX KW recessive disorder; phenotype; ss.

XX OS Mus musculus.

XX PN W09724435-A1.

XX PD 10-JUL-1997.

XX PF 30-DEC-1996; 96WO-US020785.

XX PR 29-DEC-1995; 96US-0009409P.

XX PR 29-DEC-1995; 96US-00580539.

XX PR 30-JAN-1996; 96US-00108359P.

XX PR 30-JAN-1996; 96US-00594242.

XX PR 12-APR-1996; 96US-00632175.

XX PA (DARW-) DARWIN MOLECULAR CORP.

XX PI Oshima J, Fu Y, Yu C, Mulligan J, Schellenberg GD;

XX DR WPI; 1997-363671/33.

XX PT Isolated nucleic acid molecule encoding the WRN gene product - useful for

XX PS detection and treatment of Werner's syndrome, and related diseases.

XX PS Claim 1; Fig 7; 159pp; English.

XX This sequence represents a fragment of the genomic sequence containing
 CC the coding region for the mouse WRN gene (AAX83004). The corresponding
 CC human gene (AAX83001) encodes a protein related to Werner's syndrome. The

CC products can be used for the detection and treatment of Werner's syndrome
 CC (WS), an autosomal recessive disorder with a complex phenotype, as well
 CC as related diseases

SQ Sequence 16442 BP; 4392 A; 2975 C; 3408 G; 5665 T; 0 U; 2 Other;

Query Match 14.2%; Score 323.899; DB 1; Length 16442;
 Best Local Similarity 64.1%; Pred. No. 23;
 Matches 588; Conservative 0; Mismatches 296; Indels 33; Gaps 7;

Qy 1232 TGGAAATATTATATTCATATTTTCTTGAATGTGGTAACATCTTTAGATGAAGTTT 1291

Db 28 TTGAAGATTAATCTTCTTATTTTCTAGGGCGTGTCTATCTCTTATTTGGGTTTT 87

Qy 1292 TCTCCTAGCCT-----TCTTATGCTCTGCAATTTGAAGATAGATATCTTTACATC 1341

Db 88 TTTTCTTCTGTTATTTATCTTTGAAGGCTGGATTCGTGGAGAGATAAATGTGAAT 147

Qy 1342 TGATTTTATCTAGATGCTTCTTCTCCAACTATTCTGACAGAAAGTTTCTTAAGT 1401

Db 148 TGGTATTGTCATGGAATACTTTGTTTCTCCATCTAATGGCAATTGAGAGTTTGGTTGGGT 207

Qy 1402 GCAGTAGTCTGGCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCC 1461

Db 208 ATAGTAGCCTGGCTGGCTTTGTGTTCTCTTAGGCTCTTTATAACATCTGCTAGGATC 267

Qy 1462 TTCTTACATTTGAGTTTCTTATTTGAAAGTCTGAGGTGTAATCTTAATACATCTGCCTTTA 1521

Db 268 TTCTGGCTTTCATAGTCTCTGCTGCAAGGCTCTGATATAATCTGATAGGCCCTGCCTTTA 327

Qy 1522 TATGTTAAATGGTCTTTTCTTCCATCTTTTAAATATTTCTTCTTCTTCTATACTTT 1581

Db 328 TATGTTACTT-GACITTTTCTCTTACTCTCTTTTAAATATCTATCTTTATTAGTCACT 386

Qy 1582 TAGTGATTTGATTTATGACTGTGGGAGTTTCTTTCCGGTCCAAATCTATTGGTGT 1641

Db 387 TGTGTTCTGATTAATGATGTTGGGAGGAATTTCTTTCTGCTGCTGCTATTGGAGT 446

Qy 1642 TTGTATGCTTCTGTACCTTGTAGTAGGCACTCTTCTCAAGGTTAGGAAATTTCTTT 1701

Db 447 TCTGAGGCTTCTGTATGTTCATGTGCACTCTTT---AAGTTTGGAAAGTTTCTTTC 502

Qy 1702 TTGGTTTTCTTTGAAAATAATTTTCCCTGCTTTTGACCTG-CCTTCTTCCCTTCCCTCTAT 1760

Db 503 TATTAATTTGTTGAAGATATTTGTTGGCCCTTTAAGTTGAAAATCTTCAATTTTCATCTAC 562

Qy 1761 TCC-----TTTGGTTTTGCATAGTGTCTCTGGCTTCTGAGATTTTATGCT 1807

Db 563 TCTATATCCGTGTTTGGACCTTCTCATGCTGCTGGAATTTCTGGAATGTTTAAAGT 622

Qy 1808 CTGGATTAATTTTAGACTTAACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGT 1867

Db 623 TAGGATCTTTTGCATTTTGCATTTTCTTTGATGTTGCTATGTTCTCTATGGAATC 682

Qy 1868 TTCACGTGCTGAGATCTCTCTCTATCTCTTGTATCTTGTATCTGTGAGGAGCTTGTCTGA 1927

Db 683 TTCTGCACCTGAGATTTCTCTCTCCATGCTTGTATCTGCTGCTGATGCTTTCATCTAT 742

Qy 1928 GGTTCCTG-TTGGTCTTAAATTTTCTTATTTCCAGATTTCTTCTAGTTTGGTTTGT 1986

Db 743 GGTTCAGATTTCTTCTTAGGGTTTCTATCTTAGCGTTGCTCATTTTGGGTTTCTT 802

Qy 1987 TATTAATCTATTTCCACTTTCAGGTCCTGAAATGTTTACTCATTTTCTCTCCAGTATT 2046

Db 803 TATTGTGCTACTTCGCTTTTATAGGCTACTATGTTGTTTGTGTTCAITTTCCATCCTATT 862

Qy 2047 ---TACATTTTCTAGGTTTCTTTAATGGAATTAATCTCTCTTCAAGACCTTTA 2103

Db 863 GGATGTGTTTCTCTTCTTTAAGGACTTCTACCTGTTTGGTTATTTTCTGTTT 922

Qy 2104 TGAATTTCAATAAATGTA 2120

Db 923 TTCTTTAAGGACTTTGTA 939

RESULT 25
 ABX77189
 ID ABX77189 standard; DNA; 41400 BP.
 XX AC ABX77189;
 XX DT 01-MAY-2003 (first entry)
 XX DE Mouse BAC279 clone containing alpha-1-acid glycoprotein 1 gene.
 XX KW Human; ss; transgenic; drug metabolism; behaviour; gene; mouse;
 KW pharmacokinetic assay; pharmacodynamic assay; toxicology; serum albumin;
 KW alpha-acidic glycoprotein; CYP; multidrug resistance protein; MRP;
 KW uridine diphosphoglucuronosyl transferase; UGT; cytochrome P450.
 XX OS Mus sp.
 XX WO200283897-A1.
 XX PD 24-OCT-2002.
 XX PF 18-APR-2002; 2002WO-AU000485.
 XX PR 18-APR-2001; 2001AU-00004467.
 XX PA (GENE-) GENE STREAM PTY LTD.
 XX PI Daily JW;
 XX DR WPI; 2003-093021/08.
 XX PT New transgenic non-human animal expressing a foreign polypeptide
 PT associated with drug behavior and/or metabolism, useful for studying the
 PT behavior and/or metabolism of a drug in other animals.
 XX PS Disclosure; Page 327-349; 408pp; English.
 XX CC This invention relates to a transgenic non-human animal which may be used
 CC for assessing the behaviour and/or metabolism of a drug in another animal
 CC and which expresses a foreign polypeptide associated with drug behaviour
 CC and/or metabolism. The invention also comprises a nucleic acid construct
 CC for use in producing the above transgenic non-human animal and a method
 CC of assessing the metabolism and/or behavior of a drug in an animal of
 CC interest, comprising administering a test agent to the transgenic animal
 CC and conducting analytical tests to determine drug metabolism and/or
 CC behaviour. The transgenic animal is useful in studying drug metabolism
 CC and/or behaviour in other animals. The nucleic acid construct is useful
 CC in producing the above transgenic animal and the methods are used for
 CC producing, breeding and using transgenic animals for pharmacological
 CC (e.g. pharmacokinetic or pharmacodynamic assays) and/or toxicological
 CC studies. Nucleic acid sequences used within the invention are serum
 CC albumin; alpha-acidic glycoprotein; cytochrome P450 (CYP); uridine
 CC diphosphoglucuronosyl transferase (UGT); multidrug resistance proteins
 CC and (MRP's). The present sequence represents a DNA sequence used to
 CC create a transgenic animal within the scope of the invention
 XX SQ Sequence 41400 BP; 10267 A; 10158 C; 10127 G; 10848 T; 0 U; 0 Other;
 Query Match 13.3%; Score 302.299; DB 1; Length 41400;
 Best Local Similarity 68.0%; Pred. No. 18;
 Matches 529; Conservative 0; Mismatches 222; Indels 27; Gaps 8;
 QY 1232 TGGATTATTTATTTATTTCTTCAATTTCTTGGTAAACATCTTTAGATTGAAGTTT 1291
 DB 3605 TTGAAGATTACTTTCTTCTTCTTCTAGGTGATTTTCTCTCTTGTGTGTTT 3664
 QY 1292 TCTCTAGCCTTT--TAGGTCTGCATTTGAAGATAGATATTTCTTACATCTCATTTA 1349
 DB 3665 CCATCTATTATTTCTTGTAGGCTGGATTCTGGAAGATATTGTAGATTCTGGTTTG 3724
 QY 1350 TCTTAGATGTCTTTCTTCTTCTCAACTATTGTGACAGAAAGTTTTTCTTAAGTCAGTAGT 1409

DB 3725 TCATGAATATCTTCTTTTCTCCATCTATGTAATTGAGAGTTTCTGGGTATAGTAGC 3784
 QY 1410 CTGGCTGACATCTGTAGTCTTCTGGAGTCTGTAGCACATCTGTGACGGCTTCTTACA 1469
 DB 3785 CTGGCTGGCAATTTGTCTTCTTAGGCTCTGTATGACATCTGCCAGGATCTTCTGGCT 3844
 QY 1470 TTTTGAGTTTCTATTGGAAGTCAAGTGAATCTTAATACATCTGCCCTTATATGTTAA 1529
 DB 3845 TTCTAGTCTCTGGT-GAGAAGTCAGGTGAATTTCT-TAGGTCCTGCTTATATGTTAC 3902
 QY 1530 TTGGTCTTTTCTCCCTTGCATCTTTTAATATTTCTTCTTCTTGTCTATACATTTTAGTGATT 1589
 DB 3903 TT-GACCITATTTCCCTTACTGCTTTTAATATTTCTTCTTGT-----TAGTGGTGT 3955
 QY 1590 TGATTATTATGACATCTGGGGAGTTTCTTTCCGGTCCAATCTATTGGTGTGTGATG 1649
 DB 3956 TGAATATAATGTGATGGGAGAAATTTCTTTCTCGTCCAACTATTGGAGTCTGTAGG 4015
 QY 1650 CTTCTGTACCTTGATAGGATCTCTTCTCAAGGTAGGAAATTTCTTTTGTGTTT 1709
 DB 4016 CTTCTGTATGTTTATGGCATCTCTTTT-AGTTAAGGACGTTTCTTCTATAATTT 4074
 QY 1710 TCTTGAAATATTTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCC 1755
 DB 4075 TGTGAAGATATTTACTGGCCCTTTAAGTTGGGATCGTCACTCTCTCTATACCTATTA 4134
 QY 1756 TCTATCTCTTGGTTTTCATAGTCTCTGGCTTCTGGATGTTTATGCTGGATTA 1815
 DB 4135 TCTTAGGTTTGGTTTCTCATTTGTCATGTTCCGGATGTTTGGGCTAGAGCT 4194
 QY 1816 TTTTAGACTTAACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTGTCTTCACATGC 1875
 DB 4195 TTTTGTCTTTTGTGTTTCTTTGACTGTTGTGTCATGTTTCTATGGAATCTTCTGCAC 4254
 QY 1876 CTGAGATCTCTTCTATCTTCTGATCTCTGTCAGTGAAGCTTCTCTGAGGTTCTCTG 1935
 DB 4255 CTGAGATCTCTCTTCTATCTCTTGTATTTGTTGTTGATGCTTGCATCTATGACTCTG 4314
 QY 1936 -TTGGTCTTCAATTTTCAATTTCCAGATTTCTTCAAGTTTGGTGTGTTTGTATTA 1992
 DB 4315 ATTCTTCTAGGTTTCTATTTCCAGGTTGTTCTCCCTTTGTGATTTCTTAACATA 4372
 RESULT 26
 AAS99306
 ID AAS99306 standard; DNA; 46765 BP.
 XX AC AAS99306;
 XX DT 12-MAR-2002 (first entry)
 XX DE DNA encoding aldehyde dehydrogenase 5 family, member A1 (ALDH5A1).
 XX KW Aldehyde dehydrogenase 5 family member A1; ALDH5A1;
 KW succinate-semialdehyde dehydrogenase; gene therapy; probe;
 KW antisense technology; allele specific oligonucleotide; ASO;
 KW 4-hydroxybutyric aciduria; metabolic disease; transgenic animal;
 XX chromosome 6p22; ds.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT variation /tag= a
 FT FT replace(3749,T)
 FT FT /standard name= "Single nucleotide polymorphism"
 FT FT replace(3800,G)
 FT FT /tag= b
 FT FT /standard name= "Single nucleotide polymorphism"
 FT FT replace(3829,T)
 FT FT /tag= c
 FT FT /standard name= "Single nucleotide polymorphism"
 FT FT replace(3875,G)


```

FT FT      /tag= d
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(3876,A)
FT FT      /tag= e
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(3341,T)
FT FT      /tag= f
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      4050..42765
FT FT      /tag= h
FT FT      /product= "ALDH5A1"
FT FT      /note= "Aldehyde dehydrogenase 5 family member A1.
FT FT      Specifically claimed in Claim 25"
FT FT      4050..4403
FT FT      /tag= g
FT FT      /number= 1
FT FT      replace(4059,G)
FT FT      /tag= i
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(4111,A)
FT FT      /tag= j
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(4155,C)
FT FT      /tag= k
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(4179,G)
FT FT      /tag= l
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      4404..11575
FT FT      /tag= m
FT FT      /number= 1
FT FT      replace(4430,G)
FT FT      /tag= n
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      11576..11659
FT FT      /tag= o
FT FT      /number= 2
FT FT      11660..12315
FT FT      /tag= p
FT FT      /number= 2
FT FT      12316..12486
FT FT      /tag= q
FT FT      /number= 3
FT FT      replace(12415,T)
FT FT      /tag= r
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(12422,T)
FT FT      /tag= s
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      12487..13921
FT FT      /tag= t
FT FT      /number= 3
FT FT      replace(12559,T)
FT FT      /tag= u
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(13855,G)
FT FT      /tag= v
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(13917,T)
FT FT      /tag= w
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      13922..14038
FT FT      /tag= x
FT FT      /number= 4
FT FT      replace(14021,T)
FT FT      /tag= y
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      14039..24219
FT FT      /tag= z
FT FT      /number= 4
FT FT      replace(14046,A)
FT FT      /tag= aa
FT FT      /standard_name= "Single nucleotide polymorphism"

```

```

FT FT      variation
FT FT      /tag= ab
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      24220..24363
FT FT      /tag= ac
FT FT      /number= 5
FT FT      24364..29453
FT FT      /tag= ad
FT FT      /number= 5
FT FT      24598..31819
FT FT      /tag= ae
FT FT      /number= 6
FT FT      replace(29397,T)
FT FT      /tag= af
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      29454..29597
FT FT      /tag= ag
FT FT      /number= 6
FT FT      replace(29469,A)
FT FT      /tag= ah
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      31820..31978
FT FT      /tag= ai
FT FT      /number= 7
FT FT      31979..37049
FT FT      /tag= aj
FT FT      /number= 7
FT FT      37050..37219
FT FT      /tag= ak
FT FT      /number= 8
FT FT      37220..41171
FT FT      /tag= al
FT FT      /number= 8
FT FT      41172..41230
FT FT      /tag= am
FT FT      /number= 9
FT FT      replace(41217,C)
FT FT      /tag= an
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      41231..42559
FT FT      /tag= ao
FT FT      /number= 9
FT FT      replace(41237,A)
FT FT      /tag= ap
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(41403,C)
FT FT      /tag= aq
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(42380,T)
FT FT      /tag= ar
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      42560..42765
FT FT      /tag= as
FT FT      /number= 10
FT FT      WO200190119-A2.
FT FT      XX
FT PN      29-NOV-2001.
FT PN      XX
FT PD      21-MAY-2001; 2001WO-US016558.
FT PF      XX
FT XX      19-MAY-2000; 2000US-0205849P.
FT XX      PR
FT XX      (GENA-) GENAISSANCE PHARM INC.
FT XX      PA
FT XX      Kliem SE, Koshy B, Tanguay DA;
FT PI      WPI; 2002-089912/12.
FT XX      DR
FT XX      P-PSDB; AAU73594.
FT XX      New genetic variants of human aldehyde dehydrogenase 5 family, member A1,
FT PT      ALDH5A1 gene for treating metabolic diseases and for expressing ALDH5A1
FT PT      protein useful in identifying drugs to treat 4-hydroxybutyric aciduria.

```



```
XX Human; breast specific gene; breast cancer; differential expression;
KW cytostatic; gene therapy; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200259271-A2.
XX
XX 01-AUG-2002.
XX
XX 25-JAN-2002; 2002WO-US002176.
XX
XX 25-JAN-2001; 2001US-0263757P.
XX
XX 23-APR-2001; 2001US-0286090P.
XX
XX 23-MAY-2001; 2001US-0292517P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Orr MS, Nation M, Diggins JC, Zeng W;
XX
XX WPI; 2002-674803/72.
XX
XX Diagnosing breast cancer in a patient comprises detecting the level of
XX gene expression in cell or tissue samples, where a differential gene
XX expression is indicative of breast cancer.
XX
XX Claim 1; SEQ ID NO 850; 260pp + Sequence Listing; English.
XX
XX The present invention relates to methods of diagnosing breast cancer in a
XX patient, which comprise detecting the level of expression in a tissue
XX sample of two or more genes selected from those shown in ABR09867-
XX ABR1112, where a differential expression of the genes indicates breast
XX cancer. The methods are useful in diagnosing, treating, detecting the
XX progression, and in monitoring treatment of breast cancer in patients.
XX The methods are also useful as a screening tool for agents that modulate
XX the onset or progression of breast cancer. The breast cancer genes may be
XX used as diagnostic markers for the prediction or identification of the
XX malignant state of breast tissue, for confirming the type and progression
XX of cancer, and for drug screening and assays. The present sequence is a
XX coding sequence of the invention. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub.published_pct_sequences
XX
XX Sequence 70665 BP; 22090 A; 12723 C; 13020 G; 22832 T; 0 U; 0 Other;
XX
XX Query Match 12.2%; Score 277.897; DB 1; Length 70665;
XX Best Local Similarity 52.5%; Pred. No. 17;
XX Matches 895; Conservative 0; Mismatches 726; Indels 83; Gaps 14;
XX
XX 322 TTTAACTCAGTAGTGCAGTTGTTGGTTTCCATAAGTTTGTAAAGTTTCTGTTGTTCTGT 381
XX 49114 TTTCAATCAGAGCAGGTTATTTAATTTCCATGATTTGATGGTTTGGAGTTCC--- 49058
XX
XX 382 TGTGTTGTTGTTATCTAGATTTAAAGCTGCTGGTGCAGATAGACATAGAGTATATTT 441
XX
XX 49057 TTTAGGAGTTGATTTCCAGTTTATTTCCACTGTGCTGTGAGAGAGTGCCTGATGTAATTT 48998
XX
XX 442 CAATGTCTTTATCTGTCGAGACTGTGTTTGGTTTGAATATCTATTCAATTTGGAGA 501
XX
XX 48997 CAATTTCTTAAATTTAAGGCTCATTTTATGGCCCATCACATAGTCTATCTTGGAGA 48938
XX
XX 502 GTTTCATAGGCTGTGACAAAGGTACAGTCT-TTGTGTTTGGTGAATAGTCTGTAA 560
XX
XX 48937 AAGTTCATGCTGATTTCAATCACTGATCTGTCAGTGTTCGATGAATGTTATGTAT 48878
XX
XX 561 ATATCT-----CTAGGTCCACTTGGTTTATGACATCAGTTAGTCCAGCAATT 607
XX
XX 48877 ATATATATATATATATCTGTTTAAAGTCCATTTGTTCCAAAGTTATAGTTTAAATCATTGTT 48818
XX
XX 608 TCTCTGTTTCTGTTTTTCTGTCAGATGACCTAACTGTTGAGAGAGATGGGTATTGAAGTA 667
XX
XX 48817 TCTTGTGCTCTTCTT-TCTTGATGACCTGCTAAAGCTCTCAGTGGAGTATTGAAGTC 48759
```

```
QY 668 GCCCACTATCTGTGTG-TGAGGTCAATATGTGATTTTAGCTGTAGCTGTGCTGTTTAT 726
Db 48758 CCCCACTATTATTGTGTGCTGTATCTCAATTTCTTAAGTCTATTAGTAAATGTTTAA 48699
QY 727 GAACITGGGTGACATTTGTTTGGTGCATAGACATTAAGAAATGCAATCTCTCTGGTG 786
Db 48698 ACATTTGAGATCTCCAGTGTTAGATGCAATATGTTTAGGATTCGATATTTTCCGTGCG 48639
QY 787 GATTTTCCCTTTGATGCTATGTAGTATTTCTCCCAATCTCATCTGCTTAGTTTGGGTTT 846
Db 48638 GACGGGCAATTTGCCAATATAAATGTCATCTTTGTCTCTTTAAATATTGTTGCTTTA 48579
QY 847 AAGTCTATTAG-TCAGATATTAAATGACTGTATCGGCTTGTCTCTTAGGGCCATTTGCT 905
Db 48578 AAGTTTGTTTTATCTGATATAAGAAATAGCTCCCCCGCTCACTTTTGGTGTCCATTGCA 48519
QY 906 TAGAATATCTTTTCCATCCTTTTACTCTAAGGTGATGTCATCCATGGTAGTGTGCTTT 965
Db 48518 TGAATAGCTTTTCTACCTCTTAAGTTTATGTAGTCTTATGTGTTA-GATAAGTCT 48460
QY 966 TTGGATGAGCAGTAGGATGATCTTGTGTTTCAATCCATCTGTTACCCAGTATCTTT 1025
Db 48460 CCTGAAGGCAGCAGATATTTGGTTCATGAGTTCTTATCCATCTGCACTTCTGTGTCTTT 48400
QY 1026 TTCTAGAGAAATTAAGATCAATTGAGTCATTGATGTTGAGAAATTAATCAATGACAGTGT 1085
Db 48399 TAAGTGGAGCATTTAG-----GCCATTTACATTCAGTTAGTATTGAATGAGG 48348
QY 1086 GTGATTTCTTTGTTATCTTCACCTGTGAAAGTGTGTGTGTGTGTGTGTGTGTGTG 1145
Db 48347 TACCATTGGTTAATCATGCTCT-----TTGTGCTGTGTGTGTGTGTGTGTGTGTGT 48295
QY 1146 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1205
Db 48294 TTTTTTTTCTTTTAACTGATTTTGTGTTTATAGTCTGTGTGTGTGTGTGTGTGTGTTA 48235
QY 1206 TCTCTCCCTCTTTGATTTTGGCTGGAATTAATTAATTAATTAATTAATTAATTAAT 1265
Db 48234 AAGAGTCTGTGTTGATGGTTTCCAGGATTTGTTTCAAGAT----- 48192
QY 1266 GGTAAACATCTTTAGATGGAAGTTTCTCCAGGCTTCTTTAGTGTCTGCATTTGAAGAT 1325
Db 48191 -----TTAGAGCTCTCTTTTCAATTTCTTGTAGTGTGTGTGTGTGTGTGTGTGT 48148
QY 1326 AGATATTTCTTACATCTGATTTTATCTTAGAATGCTTTTCTTCTCCAACTATTGTGACA 1385
Db 48147 GCAATTCACAGATCAATTTGTTGCTGAAATGACTGTATCTTCTCCTCATGTATGATG 48088
QY 1386 GAAAGTTTTTCTAAGTGCAGTGTGTGCTGCTGACATCTGTAGTCTCTTGGAGTCTGTAGC 1445
Db 48087 CTTAGATTCACCTGGATACAAAATTTCTGGCTGATAATGTT- -TTGCTTTGAGCGGTTGA 48030
QY 1446 ACATCTGTGCGAGGCTTCTTACATTTGAGTTTCTATTGGAAGTCAAGTGTATTTCT 1505
Db 48029 AGATAGGCGCAATCCCTTCTAGCTGTAGAGTTTCTGCTGAGATATCTGCTGTAAATCT 47970
QY 1506 AATACATCTGCTTTTATATGTTAATTTGTTCTTTTCCCTTGCATCTTTTAATTAATCTTT 1565
Db 47969 AATAGTTTACCTTTATAGTTTACCTGGT-GCTTCTGTCTCACAGCTCTTAAGATCTTT 47911
QY 1566 CTTTGTCTATACTTTTGTAGTATTGATTAATGCACTGTGGGAGTGTCTTTTCCCGT 1625
Db 47910 CTTTGTCTTAACTTTGGATAAAGTGTGAGGAAATGTTGTAGTTGAAGATCTTTTAGCAA 47851
QY 1626 CCAATCTATTGTGTGTTTGTATGCTTTCTGTACCTTTGATAGGCACTCTTCTTCAAGT 1685
Db 47850 TGAATTTCCAGGTGTTCTTTGTGCTTCTGTATTTGCAATGCTAGGTCTCTAGCAAGC 47791
QY 1686 TAGGAAATTTTCTTTTGTGTTTCTTGAAATATTTTCCCTGCTTTT-GACCTGCTTT 1744
Db 47790 CAGGCAAGTTTCTTGTATTATCCCCCAATATGTTTCCAAAGCTTTTAGAATTTGCTTA 47731
```


QY 1777 TAGTGCTCTGGCTTCTGGATGTTTATGCTCGATTAATTTAGACTTAACAATTTCTT 1836
 Db 31507 CATAAATCCAGATTTCTTGGAGGCTTGTTCATATTTCTTATTTCTTCTTGTCTT 31566

QY 1837 TGACCAAGGTATCCATTTCTTCTATCTGTTCTTCACTGCTCGATTTCTCTCTATCT 1896
 Db 31567 TGCTGGATGGGTAATTCAT-----AGACCTTAAGCTCTGAATTTCTTTATTCACCT 31620

QY 1897 CTGTATTTCTGTCAGTGAGGCTTGTCTCTGAGGTTCTCTGT-----TGGGTCTTAA 1947
 Db 31621 GTTCAGTCTATTTCTGAGACTTTCCAGAGAAATTTCAATTTCTGAAGTGTGCCAAG 31680

QY 1948 TTTTTCATTTCCAGATTTCTTCTGAGTTTGGGTTTGTATTAATTTCTATTTCCACTTT 2007
 Db 31681 TTTCTGAGTTTCTGACTGTGGTTTCTTTAAGCTATCTATTTCCATGAATATTTCTCCCT 31740

QY 2008 CAGTCTCGAATGTTTCTTACTCATTTTCT 2037
 Db 31741 TCACCTCTGTATCAATTTTGAACCTTCT 31770

RESULT 33
 ID AAA64139/c
 XX AAA64139 standard; DNA; 50000 BP.
 AC AAA64139;
 XX
 DT 20-DEC-2000 (first entry)
 XX
 DE Nucleotide sequence of a beta-tubulin antigen.
 XX
 KW Beta-tubulin antigen; inner ear protein; Meniere's disease; autoantibody;
 KW chronic ear disease; autoimmune disease; ss.
 OS Homo sapiens.
 XX
 PN WO200050593-A1.
 XX
 PD 31-AUG-2000.
 XX
 XX 25-FEB-2000; 2000WO-US004795.
 XX
 XX 25-FEB-1999; 99US-0121549P.
 XX
 PA (UYTE-) UNIV TENNESSEE RES CORP.
 XX
 PI Yoo TJ;
 XX
 DR WPI; 2000-558400/51.
 XX
 PT New beta-tubulin antigen in the membranous structure of the inner ear,
 PT reactive with antibodies of patients with Meniere's disease, for
 PT diagnosing Meniere's disease and distinguishing this disease from other
 PT autoimmune ear diseases.
 XX
 PS Claim 3; Page 51-74; 115pp; English.
 XX
 CC The present sequence encodes a beta-tubulin antigen. The protein is an
 CC antigen of the membranous structure of the inner ear protein, and is
 CC reactive with antibodies from patients having Meniere's disease.
 CC Meniere's disease is a chronic ear disease with unknown etiology. Serum
 CC from patients suffering from this disease contain autoantibodies against
 CC a 30 kDa cochlear protein antigen. The disease is believed to be an
 CC autoimmune disease. The beta-tubulin antigen is useful as a target
 CC substance in diagnosing or detecting Meniere's disease and in
 CC distinguishing this disease from other autoimmune ear diseases
 XX
 SQ Sequence 50000 BP; 16814 A; 10211 C; 9526 G; 13449 T; 0 U; 0 Other;
 Query Match 11.8%; Score 268.097; DB 1; Length 50000;
 Best Local Similarity 52.5%; Pred. NO. 21;
 Matches 957; Conservative 0; Mismatches 754; Indels 113; Gaps 18;

QY 298 TCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTACGTTGTTGGTTTCCATAAGT 357
 Db 30648 TCTTGATTTTCATTTGTTGAGCCCAATGATCATTCAGGAGCAGGTTATTTAAATTTCCATGCAT 30589

QY 358 TTGTAAGTTTCTGTTGTTTCTGTTGTTGTTGTTTATCTAGATTTAAGCTGTGGTGGT 417
 Db 30588 TTGCATGGTTTGGAGGTTCTTTTATCATTTGATTTCAAT---TTTATCCACTGTGTT 30532

QY 418 CAGATAGGACATPAGAGTATTAATTCAAATTCGTTTTTATCTGTCGAGACTTCTGTTGTTT 477
 Db 30531 CAGAGAGAGTACTTGCATATAACTTCAATTTTCTTAAATTTATTTAGACTTGTTTTGTGGC 30472

QY 478 GAAATATGTAATCAATTTTGG--ACAGTTTCTATAGGTTGCTGACAAAGAGTACAGTCT- 534
 Db 30471 CTGTCATATGGTCTATCTTGGTAAATGTTCCAGGTGCTGATGAATAGATATGATATCTG 30412

QY 535 TTGCTGTTTGGTGAATAGTCTGTAAATATCTTCTAGTCCACTTGGTTTATGACATCAGT 594
 Db 30411 CAGTTGTTGGGTAGAATGTTCTGTAAATATCTGTTAAGTCCATTTCTTTTAGGGTATAGT 30352

QY 595 TAGCTCCAGCATTTCTCTGTTTCGTTTTTGTGAGATGACCTACTGTTGAGAGAATG 654
 Db 30351 TTAAGTCCATGTTCTTCTGTTGACTTCTCTGCTTGTGATGACCTGCTAGTCTGTCAGTG 30292

QY 655 GGTATTTGAAGTAGCCCACTATCTGTGTG-TGAGGTCAATATGTAATTTTAGCTGTAGCT 713
 Db 30291 GAGTATGAGTCCCTCACTATTAATTTGTTGCTTTCTATCTCATTTCTTAGTCTAGTA 30232

QY 714 GTGCTGTTTATGAACCTTGGGTGACATTTGTTTGGTGCAATAGACATTAAGAATTCGAA 773
 Db 30231 GTAATTAATTTATACATTTGGGAGCTCTAGTGTAGGTACATATATTTTAGGATTTGA 30172

QY 774 TGFCTCTTTGGTGA-TTTTCTTTGATGCCCTATCTAGTATTTCTTCCCAATCTCATCTGC 832
 Db 30171 TATTTCTCTGTTGGCAAGTCCCTTTTATCAATTAATATGTCCTCTTTGTTGTTTAAAG 30112

QY 833 TTAGTTTGGGTTTAAAGTCTATTA-GTCAGATATTAATAAGTACTGTATCGGCTTCTCT 891
 Db 30111 CTGCTATTTGTTGAAGTTGTTTGTGCTGATATAAGAATACCTACTCTCTGCTCATTTT 30052

QY 892 TAGGGCCATTTGCTTAGATATCTT-TTCCATCTTTTACTCTAAGGTGATGCTCATCCA 950
 Db 30051 GGTGCTATTTGTCATGGAATATCTTCTTCCATCTCCTTTATGTAAGTTAATGTAATCT 29952

QY 951 T---GGTAGGTTGTCCTTTTGGATGCAGCAGTAGGATGGATCTTTGTTTTCATATCCATT 1007
 Db 29991 TATGTTGTTGGTAAGTCTCTTGAAGACACAGACATACTTGTGATGAATTTCTTATCCATT 29932

QY 1008 CTGTTACCCAGTATCTTTTCTAGAGAATTAAGATCATTTGATGCTATGATGTTGAGAT 1067
 Db 29931 CTGCCATCTGATCTTTTGTAGTGGAGCATTTAG-----GTTATTTCAATCAATGT 29880

QY 1068 TATCAATGAGCAGTGTGTTGCGAATCTTCTTATCTTGCACCTTGTGAAGTGTGTGTGTG 1127
 Db 29879 TACAATTTGAGATGTGAAGTACTTCTTATTTATATGCT-----29842

QY 1128 TGTGCTG 1187
 Db 29841 -----TAGTTGTTGCTGAAATACCTTGTGTTTCTTTTCTTTTCTTTTCTTTT 29801

QY 1188 TGTCTTGTGTGTGTGTGTCTCTCCCTCTTTTGTGATTTTGTGCTTGGAAATTTATTTAT 1247
 Db 29800 TGTATATTTTATAGGTCCTGTGAGATTTATGCTTAAGGGGTTCTATTTTGTGTCAT 29741

QY 1248 TCATATTTTCTTGAATGTGGTAAACATCTTTAGATTTGAAGTGTGTTCTCTAGCCTTCT 1307
 Db 29740 TTTG-----ATGATTTGTTTCAAAATTTAGAGCTACTTTTATCA 29702

QY 1308 AGTCTGCATTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAATGCTTTCTT 1367
 Db 29701 GTGCTTGTAGTGTGCTTGGTAGTGGCTAATTCAGCATTTGCTTGAAAGAACTATCTT 29642

```
QY 1368 TCTCCAACTATTGTGACAGAAAGTTTTTCTAAGTCAGTAGTCTGGCCTGACATCTGTAG 1427
|||
Db 29641 TCT--TTCAATTATAAGCTTAGTTTCACTGGACACAAATCTTGGCTGATATGTGTT 29584
|||
QY 1428 TCTCTGGAGTCGTGAGCAGATCTGTGAGGCGCTTCTTACATTTGAGTTTCTTATTGGA 1487
|||
Db 29583 TCTTTAAGGAGGCTAAAGATAGAACCCCAATCCCTTCTAGCTGTAGGTTTCTGCT-GA 29525
|||
QY 1488 AAGATCAGGTGAATCTTAATACATCTGCCCTTATATCTTAATTTGGTCTTTTCCCTTG 1547
|||
Db 29524 GAAATCTCTGTTAATCTGATAAGTTTCCCTTTATAGGTTACCTGAT-GCTTTTGTCTCA 29466
|||
QY 1548 CATCTTTTAATATCTTCTTTTGTCTATATCTTTTGTAGTTGATTTGATTTATGCACTGTG 1607
|||
Db 29465 CAGCTCTTAAGATCTTCTTCTTGTCTGACTTTAGATAAACCCTGATGACTAAGTGCCTAC 29406
|||
QY 1608 GGGAGTTCTTTTCCGGTCCCATCTAATTTGGTGTGTTGATGCTTCTTGACCTGTATAG 1667
|||
Db 29405 ATGATATCTTTTGTGATGAATTTTCCAGGTGTTCTTTTCACCTTCTTGTATTTGGGTAT 29346
|||
QY 1668 GCATCTCTTCTCAAGGTAGGAATTTTCTTTTGGTTTCTTGTGAATAATTTTCCC 1727
|||
Db 29345 CTAGATCTCTAGCAGGCCAGGAAGTTTCTTGAATTTTCTTCACTAATGTTTCCCA 29286
|||
QY 1728 TGCCTTTGACCTGCTTCTTCCCTTCTCTATCTTCTTCTTCTTCTTCTTCTTCTTCT 1776
|||
Db 29285 AACCTTTAGATATCTTCTTCTTCTTGGGAACACTAAATTAATCTTAGAATCCATCTTAA 29226
|||
QY 1777 TAGTCTCTCGCTTCCTGAGATGTTTATGCTGATTAATTTAGATTAACATTTCTT 1836
|||
Db 29225 CATATCTCAACTTCTCTGGAGCTTGTCTTCAATTTTAAATTTCTGTTTCTTGTGACTT 29166
|||
QY 1837 TGACCAAGGTATCCATTTCTTCTATCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 1896
|||
Db 29165 GTCAGATGGGTAAATTT-GAAGCCTTGTCTTCAAGCTCTGAAAGTTCTTCTTCTTACTT 29107
|||
QY 1897 CTGTATTTCTCAGTAGGCTTCTCTGAGTTCCTGTGCTTCTTCTTCTTCTTCTTCT 1956
|||
Db 29106 ATTCAATATATGCTGAGACCTTCCAGTGTATTTTGCATTTCTGTGTCTCTTCAAT 29047
|||
QY 1957 TTCCAGATTTCTTCCAGTTTGGGTTTGTGTTTATTAATCTAATTTTCCACTTTCAGGCTG 2016
|||
Db 29046 TCCAGAGTTT---GATGCTTTTATTTATGCTATTTATTTCCCTGT----- 29001
|||
QY 2017 AAATGTTTACTCATTTTCTCCAGATTTTACATTTTCAAGATTTTCTTCTTCTTCTTCT 2076
|||
Db 29000 AGATTTTCCATTCATATCTGTAACATTTTAAATTTCTTGTAGTTGGTATTCACATTT 28941
|||
QY 2077 ATTCAATTTCTTCCAGGACCTT 2100
|||
Db 28940 CTCTTCTCCCTTGTAGTAGCTT 28917
|||
```

RESULT 34

ABK83575/c

ID ABK83575 standard; cDNA; 136284 BP.

XX ABK83575;

XX AC ABK83575;

XX 14-AUG-2002 (first entry)

XX Human cDNA differentially expressed in granulocytic cells #146.

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;

XX viral infection; parasitic infection; protozoal infection;

XX fungal infection; sterile inflammatory disease; psoriasis;

XX rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;

XX cardiac reperfusion injury; renal reperfusion injury; ARDS;

XX adult respiratory distress syndrome; inflammatory bowel disease;

XX Crohn's disease; ulcerative colitis; periodontal disease;

XX granulocyte activation; chronic inflammation; allergy.

XX Homo sapiens.

OS

XX W0200228999-A2.

XX 11-APR-2002.

XX 03-OCT-2001; 2001WO-US030821.

XX 03-OCT-2000; 2000US-0237189P.

XX (GENE-) GENE LOGIC INC.

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

XX WPI; 2002-435328/46.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity.

Claim 1; SEQ ID NO 146; 114bp; English.

The invention relates to detecting (M1) granulocyte (GC) activation (GCA), by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GCA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 136284 BP; 41693 A; 28668 C; 28000 G; 37923 T; 0 U; 0 Other;

Query Match 11.7%; Score 267.297; DB 1; Length 136284;

Best Local Similarity 51.5%; Pred. No. 12;

Matches 915; Conservative 0; Mismatches 777; Indels 83; Gaps 14;

QY 298 TCTTGAATTTCTATCTTGGCTCATTTTAACTCAGTAGTGTGTTGGTTCCATAAGT 357

Db 83169 TCTTGAATTTCTGTTTGGACCAATCTCTCGGAGCAGAGTTATTTATTTCCATGAT 83110

QY 358 TTCTAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 417

Db 83109 TTGCATGTTTTCAGAGTTCCTTTTGGAGTTGATTTCCAGTTGTATTCACATGTTGTT-T 83051

QY 418 CAGATAGGACATAGAGTATTTATTTCAATTTGTTTATCTCTCGAGACTTGTGTTT 477

Db 83050 GAGAGAGGCTTAC---ATAATTTCAATTTTCTTAAATTTATTAGAGTTGTTTATGGC 82994
 QY 478 GAAATATGATATCAATTTTGGAG--AGTTTCATAGGTTGCTGACAAAGAGTACAGTCT- 534
 Db 82993 CTATCATATGCTCTATCTTGAGAGAGGTTCTATGACATGTTTAAATAGAAATGATATTCTG 82934
 QY 535 TTGTTGTTTGTGAAATAGTCTGTAAATATCTCTAGTCCACITGGTTTATGACATCAGT 594
 Db 82933 CAGTTGTTGATGAAAGTCTCTGATATATCTGTTAAGTCAATTTGTTCCAAAGTATAGT 82874
 QY 595 TAGCTCCAGCAATTTCTCTGTTTCTGTTTGTGAGATGACCTTAACCTGTTGGAGAGATG 654
 Db 82873 TTAATCCATTTGTTTCTGTTGCTCTCTGTTGATGACCTGCTAGTCTGTCAGTG 82814
 QY 655 GGGTATTGAAGTAGCCCACTATCTGTG-TGAGTCAATATGATGATTTTAGCTGTAGCT 713
 Db 82813 GTGTATTGAAGTCCCACTATTTATTTGGTTGCTGTCTATCTCAATTTCTTCAGTCTATTA 82754
 QY 714 GTGCTTGTTTTATGAACCTTGGTGACATTTGTTTGGTGCATGACATTAAGAAATGCAA 773
 Db 82753 GTAACGTTTATAAATTTGGAGCTCCAGTGTAGGTGCATATATGTTAGGACTGTGA 82694
 QY 774 TGTCTCTTGTGGA-TTTTCTTTGATGCCATGATAGTATTCTTCCCAATCTCATCTGC 832
 Db 82693 TATTTTCTGTGGACAAGGCTTTTACATTAATAATGCTTCTTGTCTCTTTAA 82634
 QY 833 TAGTTTGGGTTTAAAGCT-ATTAGTCAGATATTAAATGACTGTATCGGCTTGTCTCT 891
 Db 82633 CTGCTGTTGATTAAGTTTGTGTTGCTGTATTAAGAGTAGTACTCTCTGCTCACITTT 82574
 QY 892 TAGGCCAATTTGCTTAGAATATC-TTTTCCATCCTTTTACTCTTAAGGTGATGCTATCCA 950
 Db 82573 GGTGTCCATTTGCAATGAATGCTTTTCCACCCCTTTACTTTAAGTTTATGTCAGTCT 82514
 QY 951 TGGTAGGT-TGTCCTTTTGGATGAGCAGTAGGATGATCTGTTTTCATATCATCTCT 1009
 Db 82513 TACATGTTAGGTGAGTCTGAAGGACGACGATATTGTTGGTGAGTTTATGCAATCT 82454
 QY 1010 GTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGTGAGAAATTA 1069
 Db 82453 GCAGTTCTGTATCTTTAAATGGAGCAATTTAGTTAATTTACATCAATGTTAGTATTA 82394
 QY 1070 TCAATGACAGTGTGTTGGATCTTGTATCTTGCACCTGTGAAGTGTGCTGTGCTG 1129
 Db 82393 ATGGAGGTGACTGTGCAATGCTATTTGTTGCTGTGCTGTGCTGTGCTGTGCTGTG 82334
 QY 1130 TGT 1189
 Db 82333 GTTTTGTCTTTTAACTGTATTTTGTGTTTATAGTCTGTG----- 82292
 QY 1190 TCTGT 1249
 Db 82291 -----TGATTTATGCTTTTAAAGAGGTATGTATG 82262
 QY 1250 ATATTTCTTGAATGTGGTAACATCTTTTAGATTCAGAGTTTCTCTAGCCTCTTTTAG 1309
 Db 82261 ATGTGTTTCCAGGATTTGTTCAACATTTAGAGTTCTTTTATGTTCTTTGAGTCT 82202
 QY 1310 GTCTGCATTTGAAGATAGATATCTTTTATCATCTGATTTTATCTTAGAATGCTTTCTTTC 1369
 Db 82201 GGTGTTGGTAGTGGCAATTTCTCTGACATGTATGTGTTTGTCTGAAAAGAGACTGTATCTT 82142
 QY 1370 TCCAACTATTGTGACAGAAAGTTTTTCTAAGTGCAAGTAGTCTGGCTCAGATCTGTAGTC 1429
 Db 82141 TCCCTTCATATGATGATCTTAGTTTGTCTGATGATGATGATGATGATGATGATGATG 82082
 QY 1430 TCTTGGAGTCTGTAGCACATCTGTGAGGGCTTCTTATCATTTTGAAGTTTCTATGGAAA 1489
 Db 82081 TTTGAGGAGGCTGAAGATAGGGCCCAATCCCTTCCAGCTGTGAAGGTTTCTGTCT-GAGA 82023
 QY 1490 AGTCAGGTGATTTCTAATACATCTGCTTTATATGTTAATGTTTATGTTTCTTCTTCC 1549
 Db 82022 AATCTGCTGTTAATCTGATAGGTTTCTTCTTATAGGTACTTGGT-GCTTCTGCTCAGA 81964

QY 1550 TCTTTTAAATTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1609
 Db 81963 GCTCTTAGACATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 81904
 QY 1610 GAGTTTCTTTTCCGGTCCAACTATTTGGTGTGTTTGTATGCTTTGTGTACCTTGATAGGC 1669
 Db 81903 TATGATCTTTTGTGATGAAATTTCCAGGTGTGTTCTTTTGTCTTGTGTTGTGTTGTCT 81844
 QY 1670 ATCTCTTCTCAAGTGTAGGAAATTTCTTTTGTGTTTCTTTGAAATATTTTCCCTG 1729
 Db 81843 AGTCTCTAGCAAGCTGGGAATTTCTTTGATTTATTTCCCAAGTATGTTTCCCAAG 81784
 QY 1730 CTTTGACCTGCTTCTTCCCT-----TCTCTATTTCTTTGGTTTTTGCATAG 1779
 Db 81783 TTTTGAATTTCTTCTTCTTCTCAGGAACACCAATTTATTTCTAGTTTGGTCAATTAAGCT 81724
 QY 1780 TGTCTCTGCTTCTTGGATGTTTATGCTGGATTTATTTAGACTTAACATTTCTTTGA 1839
 Db 81723 AATCTCAGACCTCTTAGATGCTTTGTCGATTTCTTATCTTTTCTTTGCTGTAT 81664
 QY 1840 CCAAGTATCCATTTCTTCTATCTTGTCTCTCACTGCTGAGATTTCTCTCTATCTTT 1899
 Db 81663 TGGATTTGGGTTAATTTCAAGACCTTGTCTTTGAGCTCTGAATTTCTTCTTCTTCTT 81604
 QY 1900 GTATTTCTGCTAGGCTTGTCTGAGGTTCTTGTGGTCTTAAATTTTCTTCTTCTT 1959
 Db 81603 CAGTTCTATTTGCTGAGACTTTCCAGAGCATTTTGCATTTCTGTAAGTATGCTCAATGTT 81544
 QY 1960 CAGATTTCTTCTCAGTTTGGGTTTGTGTTTAAATTT-----CTATTTTCCATTT 2008
 Db 81543 CCTGAATTTTGAATGCTTTTCTTTCTTAAGCTATTTATTTCTTGAATTTCTTCTT 81484
 QY 2009 AGGTCTCGAAATGTTTTACTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2043
 Db 81483 ACTTCTGATCTTTTCTTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 81449

RESULT 35

AAL0509/c

ID AAL0509 standard; DNA; 32167 BP.

XX AAL0509;

XX AC

XX DX

XX 21-NOV-2001 (first entry)

XX DE

XX DE

XX KW

XX KM

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

Human reproductive system related antigen DNA SEQ ID NO: 8197.

Human; reproductive system related antigen; reproductive system disorder; cancer; gene therapy; ds.

Homo sapiens.

WO200155320-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US001339.

31-JAN-2000; 2000US-0179065P.

04-FEB-2000; 2000US-0180628P.

24-FEB-2000; 2000US-0184664P.

02-MAR-2000; 2000US-0186350P.

16-MAR-2000; 2000US-018974P.

17-MAR-2000; 2000US-0190076P.

18-APR-2000; 2000US-0198123P.

19-MAY-2000; 2000US-0205515P.

07-JUN-2000; 2000US-0209467P.

28-JUN-2000; 2000US-0214886P.

30-JUN-2000; 2000US-0215135P.

07-JUL-2000; 2000US-0216647P.

07-JUL-2000; 2000US-021680P.

11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217456P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225288P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232081P.
PR 08-SEP-2000; 2000US-0232082P.
PR 12-SEP-2000; 2000US-0231958P.
PR 14-SEP-2000; 2000US-0232377P.
PR 14-SEP-2000; 2000US-0232388P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 14-SEP-2000; 2000US-0233085P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235836P.
PR 27-SEP-2000; 2000US-0235837P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-02444517P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465570/50.
XX
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
XX used in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; SEQ ID NO 8197; 1297pp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a genomic sequence encoding a
XX protein of the invention
XX
XX Sequence 32167 BP; 10519 A; 6697 C; 6488 G; 8463 T; 0 U; 0 Other;
Query Match 11.3%; Score 256.797; DB 1; Length 32167;
Best Local Similarity 51.1%; Pred. No. 30;
Matches 891; Conservative 0; Mismatches 772; Indels 82; Gaps 13;

QY 290 CTTCTATTCTTGATTCTCTATCTCTGGCTCATTTTAACTCAGTAGTGAGTTGGTTT 349
Db 17941 CATTTCTCTCTGTGATTTTGGTTTGTGACCAATGCTCATTCAGGAGCAGGTTATTTAAATTT 17882
QY 350 CCATAAGTTTCTGAAGTTTCTCTGTTCTGTTCTGTTGTTGTTTATCTAGATTAAAGCT 409
Db 17881 CCCGTATTTCGATGGTTTTCGAAGTTCCT---TTGAGGTTAATTTCCAGTTTATTTCC 17826
QY 410 GTGGTGTGCATAGGACATAGAGTATATTTTCAATTTGCTCTTTTATCTGCGAGACTTGC 459
Db 17825 ACTGTGGTCTGAGACAGTCTGAGTATATAATTTCAATTTCTTAAATTTATTTGAAGTTTCAT 17766
QY 470 TTTGTTTTGAATATGATTTCAATTTTGG-----AGATTTTCATAGGCTGCTGACAA 521
Db 17765 TTTATGGCCCAATGCTGCTCTATCTTGGTGAAGTTAGAAAGTTATGCGCTGTTGAATA 17706
QY 522 GAAAGTACAGTCT-TTGTGTTTTGGTGAATAGTCTGTGAATATCTCTAGTCCACTTTGG 580
Db 17705 GAATGTGATTTCTGCTGCTGTTGGATGAATGTTCTGTATATATCTGTAAAGTCCATTG 17646
QY 581 TTTATGACATCAGTTAGTCTCCAGCATTTCTCTGTTTCTGTTTTTGTGAGATGACCTTAAC 640
Db 17645 TTCAAAGTATAGTTTAAATCCATGTTTCTTGTGTGACITTTCTGCTTGTATGACCTGTC 17586
QY 641 TGTGGAGAGTAAGGTTATTAAGTATAGCCCACTATCTGTGTG-TGAGGTCAATATATGTA 699
Db 17585 TAGTGTCTGAGTGGAGTATTAAGTCCCCACTGTTATTGTTGTTGCTATCTATTATT 17526
QY 700 TTTTAGCTGTAGTGTGCTGTTTATGAACTCTGGGTGACATTTGTTTGGTGGCATAGAC 759
Db 17525 TCTTAGGTCTAGTACAAATCTTTATATAATTTGGAGCTCCAGTGTAGGTGCTATAT 17466
QY 760 ATTAAGAAATGCAATGCTCTCTGGTGAATTT-CCTTTATGATGCTATGATGATTTCTTC 818
Db 17465 ATTTAGGATGATATTTTCTGCTGACAAAGCCTTTACATTTCTATCTCTCTG 17406
QY 819 CCAATCTCACTGCTTAGTTTGGTTTAACTATATTA-GTCAGATATTAATGACTGT 877
Db 17405 TTTGTCTCTTTTAAACCAGTGTGTTTAAAGTTTGTGTTGTTCTGATATAGGAATAGCTAC 17346
QY 878 ATCGCTTGTCTTTAGGCAATTTGCTTGAATATCTTTTCCATCTTTTACTCTAAGG 937
Db 17345 CTCGTCTCACTTTTCAATGCTCTTTTCAATGAATAGCCCTTTTACCTGTTTACTTTAAG 17286
QY 938 TGATGTCTATCCATGTTAGTTGTCTTTTGGATGACGACATAGGATGGAATCTGTTTT 997
Db 17285 TTTATGTGAGTCTTATGTGCTAGTGTCTCTCTGAAGGACGAGTTGGTTGGTGAGTT 17226
QY 998 CATATCCATTTCTGTACCCAGTATCTTTTCTAGAGAAATAGATCATTTGAGTCAATGA 1057
Db 17225 CTTATCCATTTCTGAGTCTCTATCTTTTAAAGTGAGCAATTTAGCCATTTACATTCAT 17166
QY 1058 TGTTCAGAAATATCAATGACGAGTGTGTTGGAATCTTTGTTATCTTGCACTTTGTGAAGTG 1117
Db 17165 GTTAGTATGAATCTGAGGTACTGTTGCTTTTCAATGCTCTTTGTTGCTGTGAATTT 17106
QY 1118 TGT 1177
Db 17105 TTGTTTTTGT 17046
QY 1178 TCTGT 1237
Db 17045 ATACTGTAAGAGGTGTGTT-----AT 17023
QY 1238 TATTATATTTATTTCTTTGTAATGTTGGGTAACTCTTTTATGATTTGAAAGTTTCTCTCT 1297
Db 17022 GATGTGTTTGCAGGATTTGTTTTCAGATTTAGAGCTTTTATAGAGCTCTTTGTAGTGTGT 16963
QY 1298 AGCCTCTTTTAGGCTGCAATTTGAAGATAGATATTTCTTACATCTGATTTTATCTTGAAG 1357
Db 16962 AACGTGGTAATGG-----CGAATTTCTCTAGCAATTTGTTTGTCTGAAAA 16919
QY 1358 TGTCTTTCTCCAACTATTTGTGACAGAAAGTTTTTCTTAAGTGCAGTAGTCTGGCCTG 1417

Db 16918 CCACTGTATCTTTTCTTCTCATATATGATGCTAGTTTGTCTGGATACAAAATCTTGGCTT 16859
QY 1418 ACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTCTGAGGGCCCTTCTTACATTTTGAGT 1477
Db 16858 ACATCTGTAGTCTCTTGGAGTCTGTAGAACATCTGAGGGCCCAATCTCTTCTAGCTTTAGGGT 16799
QY 1478 TTCTATTGGAAGTCAAGTCAAGTCAATCTTAATACATCTGCTTTATATGTTAATTTGCTT 1537
Db 16798 TTCTGCTG---AGAAATCTGTTAATCTAATAGGTTTTCCTTTGTAGTCCAGGCTGT- 16744
QY 1538 TTTTCCCTTGCATCTTTTAAATATCTTTCTTTTGTCTATACCTTTTGTAGTATTTATTT 1597
Db 16743 TTCTGTCTCACAGCTCTTAAAGATTTCTTCTTGGTCTTAACCTTTGGATAAACCTGATGACA 16684
QY 1598 ATGCACTGTGGGAGTGTCTTTTCCGGTCCCAATCTATTTGGTGTCTTTGTATGCTTCTTGT 1657
Db 16683 ATGTGCTAGGCAAAAGATCTTTTATGTGATGAATTTCCAGGTTCTTTTGTGCTTCATGT 16624
QY 1658 ACCTTGATAGGCACTCTCTTTCTCAAGGTTAGGAAATTTTCTTTTGTGTTTCTTCTGAAA 1717
Db 16623 A--TTTGCATGCTAGGCTAGCAAGGCTTGAAGTTTTCCTCAATTTATTCACCAATA 16566
QY 1718 ATATTTTCCCTGCTTTTGTACCTGCTTCTTCCCTC-----TTCCTCTATTTCTTTGGTTT 1772
Db 16565 TGTTTTCAAGTGTTTAGAAATTTGCTTCTTCTCAGGAACATCAAGTATTTCTTAGGTTT 16506
QY 1773 TGCATAGTCTCTGCTGCTTCTTGGATGTTTATGCTGGATTTATTTAGCTTATTTAGACTTAACATTT 1832
Db 16505 GTCTTTTAACTAGACTTCTTGGAGGCTTGTTCATATTTCTTATTTTCTTTTCTTTG 16446
QY 1833 TCTTTGACCAAGTATCCATTTCTTATCTTGTCTTCACTGCCGAGAGATCTCTCTTCT 1892
Db 16445 TCTTTATTTATTTGGGTTAA-ITCAAAGACCTTGACTTGGAGCTCTGAAATTTCTTCTTCT 16387
QY 1893 ATCTCTTGTATTTCTGAGTGGGCTTCTCTGAGGTTCCCTGAGGTTCCCTGTTGGTTCTTAATTTT 1952
Db 16386 ACTTGTTCATTTCTATGCTGAGACTTCCAGAGCAATTTAGCAATTTCTAAAGTGTGTC 16327
QY 1953 TCATTTTCCAGATTTCTTCCAGTTTGGGTTTGTTTTATTAATTTCTATTTCCACTTTCCAGT 2012
Db 16326 AAGTTTCTCGAATTTTGTCTTGTTTTTCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 16267
QY 2013 CCTGA 2017
Db 16266 ACAGA 16262

RESULT 36
AAV42558
ID AAV42558 standard; DNA; 10409 BP.
XX AAV42558;
XX AC
XX XX
DT 09-NOV-1998 (first entry)
XX Mouse dectin-2 gene.
DE Dectin-2; dendritic cell; c-type lectin; mouse; immunity; adjuvant;
KW allergy; autoimmune disease; gene therapy; vaccine; diagnosis;
XX drug screening; ss.
XX Mus sp.
XX
FH Key Location/Qualifiers
FT promoter 1..146
FT /tag= a
FT /note= "(Claim 136)"
FT exon 147..171
FT /tag= b
FT /number= 1
FT intron 172..632
FT /tag= c

(CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.

Montoliu Jose L. Giraldo Carbaio P. Busturia Jimeno AM:

WPT: 2003-112151/10.

New DNA segment with insulating activity, useful for reducing the variation of transgene expression, in animal cells, caused by integration at different genomic locations.

Claim 4; Page 54-55; 57pp; Spanish.

The invention relates to a 3711 bp region of the mouse tyrosinase gene referred to as the locus control region (LCR; ABV76777), and its functional fragments and analogues (ABV76778-ABV76781). The tyrosinase gene from which the LCR is derived is located on chromosome 7 and plays a role in pigmentation. A portion of the full-length LCR sequence, designated HS (ABV76779) was identified as a DNaseI hypersensitive region that was found to be deleted in the mutant tyrosinase allele chinchilla-mottled, which confers a variegated pigmentation phenotype. The full-length LCR, or its functional fragments or analogues may be incorporated into animal cell expression constructs to protect against the effects of transgene integration at different locations in the genome. The LCR-derived sequence reduces the variability of transgene expression associated with integration at different chromosomal locations, and ensures optimal expression (as regards time, place and quantity), regardless of the site of integration. The present sequence represents the mouse tyrosinase gene locus control region X. This region corresponds to bases 222-2617 of the full-length LCR, and retains its protective effect against variable transgene expression. Note: The present sequence is not shown as an isolated sequence in the specification, but is obtained from the full-length LCR sequence shown on page 54-55 and the information given in claim 4.

sequence 2396 BP: 625 A; 381 C; 413 G; 977 T; 0 U; 0 Other;

```
Query Match      11.0%; Score 249.899; DB 1; Length 2396;
Best Local Similarity 62.2%; Pred. No. 1.1e+02;
Matches 574; Conservative 0; Mismatches 296; Indels 53; Gaps 12;
```

292 TCTATTCTTGATTCTATCTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGGTTTCC 351

396 TTTCTTTCCTTTATTTCTGCGTTTGACCTCABAGTCATCCAGTAGGGAGTTGTGCAGCTTTC

455

252 ATAGCTTCTAGACCTTCTGTTGTTATCTAGATTAAAGCTGT 411

ATCCGCTTTTCTCCTA 489

[illegible]

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

.....

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

609 CTTTACCCTCCATTATTCCTTAGTCTTAG-CTTTGGCTAGATGCACTGTCTATTCTAGTG / 273

651 AATGGGGTATGAAGTAGCCCACTAATCTGTGTGTGAG-GTCAATATGIGATTATAGCIGI 709

6 725 AGTAGGATATGAGGTCCTCCCACTATTAATGTATAGATTCAAATTTGTGATTTGAGCAAT 784

710 AGCTGTGCTGTTTATGAACCTGGGTGACATTGTGTTGGTCATAGACATTAAAGAAATT 769

b
785 AGTAATGTTTCACTT-CCAAATGTGTGTACCCCTTGCTTTGGGGACATAGATGTTAAGAACT 843

XX 18-MAY-2001; 2001ES-00001133.
XX (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX Montoliu Jose L, Giraldo Carballo P, Busturia Jimeno AM;
XX WPI; 2003-112151/10.
XX New DNA segment with insulating activity, useful for reducing the
PT variation of transgene expression, in animal cells, caused by integration
PT at different genomic locations.
XX Claim 4; Page; 57pp; Spanish.
XX The invention relates to a 3711 bp region of the mouse tyrosinase gene
CC referred to as the locus control region (LCR; ABV76777), and its
CC functional fragments and analogues (ABV76778-ABV76781). The tyrosinase
CC gene from which the LCR is derived is located on chromosome 7 and plays a
CC role in pigmentation. A portion of the full-length LCR sequence,
CC designated HS (ABV76779) was identified as a DNaseI hypersensitive region
CC that was found to be deleted in the mutant tyrosinase allele chinchilla-
CC mottled, which confers a variegated pigmentation phenotype. The full-
CC length LCR, or its functional fragments or analogues may be incorporated
CC into animal cell expression constructs to protect against the effects of
CC transgene integration at different locations in the genome. The LCR-
CC derived sequence reduces the variability of transgene expression
CC associated with integration at different chromosomal locations, and
CC ensures optimal expression (as regards time, place and quantity),
CC regardless of the site of integration. The present sequence represents a
CC mouse tyrosinase gene mutant LCR designated LCRmut which retains a
CC protective effect against variable transgene expression. This sequence is
CC identical to the full-length LCR, apart from deletion of a 62 bp enhancer
CC (denoted as Box AB in the specification) between bases 2622 and 2683 of
CC the full-length LCR; however, this enhancer is not required for
CC protection against transgene expression variability. Note: The present
CC sequence is not shown in the specification, but is derived from the
CC sequence of the full-length LCR shown on pages 54-55 (ABV76777) and the
CC information given in claim 4.
XX
SQ Sequence 3649 BP; 1026 A; 581 C; 635 G; 1407 T; 0 U; 0 Other;
Query Match 11.0%; Score 249.899; DB 1; Length 3649;
Best Local Similarity 62.2%; Pred. No. 89;
Matches 574; Conservative 0; Mismatches 296; Indels 53; Gaps 12;
QY 292 TCTATTCCTGATTTCTATCTGGCTCATTTTAACTCAGTACGACGTGTTGGTTCC 351
DB 617 TTCTTTCTTTATTTCTGCTTGACCTCAAGTCATCCAGTAGGAGTTGCGAGCTTC 676
QY 352 ATAACTTTGTAACTTTCTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
DB 677 ATGAGTTGTAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720
QY 412 GGTGTGATAGGACATAGATATATTTCAATGCTTTTATCTGTCGACACTGCTT 471
DB 721 TGTTCTCTTAATAAGACACAAAGCATATTTCAACTTCTTATATCTGTTGAGGCTTATTT 780
QY 472 TCTTTTGAATAATGATTAATCAATTTTGGAG-AGTTTCATAGGCTGCTGCAAGAAAGGTACA 530
DB 781 TGTGAAACATATATGATCAATTTTGGAGAAAGTTTAGGATTGCTGAGAAAGGTGCC 840
QY 531 GTCCTTTGTTTGGTGAATAATGCTGTAATAATCTCTAGTCCACTTTGGTTTATGACAT 590
DB 841 TTTCTTTTGTGTTGGTGAATGTTCTTAAAAA-----ATATATATTAGGT 886
QY 591 CAGTTAGCTCCAGCAATTTCTGTTCTGTTTCTGTTTCTGTTGAGATGACCAACTGTTGGAGAG 650
DB 887 CTTTTFACCTCCATATTTTCTTAGTTAG-TTTTGTCTAGATGACATGCTTATTTCATGAG 945
QY 651 AATGGGGTATTGAAGTAGCCCACTATCTGTTGTTGAG-GTCAATATGTTGATTTTAGTGT 709
DB 946 ACTAGGATATGAAGCTCTCCCACTATTAAATGATTAGATTCAATTTGTTGATTGGCAATT 1005

QY 710 AGCTGTGCTGTTTATGAACCTTGGGTGACATTTGTTGGTGACATAGACATTAAGAATT 769
DB 1006 AGTAATGTTTCACTT-CCAAATGTTTACCCCTTGCTTTGGGGACATAGATGTTAAGAACT 1064
QY 770 GCAATGCTCTCTGG-TGGATTTTCCCTTTGATCCCTATGATGATTTCTTCCCAATCTCAT 828
DB 1065 GAAGTGTATCATGGCTAAATTTTCTTTGATGAGTATGAATGTCATTTTCCATCTCTA 1124
QY 829 CTGCTTAGTTTGGGTTTAAAGTCT-ATTAGTCAGATATTAATAATGACTGTATCGGTTGC 887
DB 1125 TTGACTAAATTTCTGTTAAAGTGTGATTTGCTAGATATTAGATAGCTAAACAGCTTGC 1184
QY 888 TTCTTAGGGCCATTGCTTAGAATATCTTTTCCATCTTTTCTCTAAAGTGTATGCTAT 947
DB 1185 TTCTTAGTGCATCTGCTTGGAAATATTTTTCACCCCTTAACCTGAGGTATTTGCTAT 1244
QY 948 C--CATGCTAGGTTGCTTTTTTGGATGAGAGTAGGATGATCTTTTTCATATCCA 1005
DB 1245 CTTGATGTTGAGGCTATTTTGTACTCAGTGAATGATGATTTCTTTTGGATCCA 1304
QY 1006 TTCTTTTACCAGTATCTTTTCTGAGAAATTAAGATCATTTGAGTCATGATGTTGAGA 1065
DB 1305 CTCTGTTGCTGCTGTTGTTT-----ATTGAGTAATGAGTCCCAITTTACTTTGAGA 1355
QY 1066 ATTATCAATGAGCAGTGTGTTGCTGATTTCTGTTTATCTTGCACCTGTGAGGTGTGTTG 1125
DB 1356 AATATTAAATGATCAATGATTTCTTCTGTTATTG-----GTGGTGGGTGATAG 1410
QY 1126 TGT 1184
DB 1411 TAGTGATAGTTGT 1470
QY 1185 CTGTGCTTGT 1207
DB 1471 CTGTATCATGATGCTATATGC 1493
RESULT 39
ABV76777
ID ABV76777 standard; DNA; 3711 BP.
XX AC ABV76777;
XX AC
XX DT 07-MAR-2003 (first entry)
XX DE Mouse tyrosinase gene full-length locus control region (LCR).
XX KW Mouse; tyrosinase; locus control region; LCR; chromosome 7;
KW transgene integration; transgene expression variability; protection;
KW chromosomal integration; transgenic animal; gene; transgenic; ds.
XX OS Mus sp.
XX Key Location/Qualifiers
FH misc_signal 222..2617
FT /tag= a
FT /note= "Locus control region X (ABV76780)"
FT misc_signal 1583..3711
FT /tag= c
FT /note= "Locus control region HS (ABV76779). This region
FT is hypersensitive to DNaseI"
FT misc_signal 1583..2617
FT /tag= b
FT /note= "Locus control region (ABV76781)"
FT mutation replace(2621..2684, GG)
FT /tag= d
FT /note= "Deletion of these 62 bp generates the locus
FT control region LCRmut (ABV76778)"
FT enhancer 2622..2683
FT /tag= e
FT /note= "Denoted as Box AB in the specification"
XX

PN WO200295031-A1.
 XX 28-NOV-2002.
 PD 10-MAY-2002; 2002WO-ES000221.
 XX 18-MAY-2001; 2001ES-00001133.
 PR (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
 XX Montoliu Jose L, Giraldo Carballo P, Busturia Jimeno AM;
 XX WPI; 2003-112151/10.
 DR New DNA segment with insulating activity, useful for reducing the
 PT variation of transgene expression, in animal cells, caused by integration
 PT at different genomic locations.
 XX Claim 1; Page 54-55; 57pp; Spanish.
 XX The invention relates to a 3711 bp region of the mouse tyrosinase gene
 CC referred to as the locus control region (LCR; ABV76777), and its
 CC functional fragments and analogues (ABV76778-ABV76781). The tyrosinase
 CC gene from which the LCR is derived is located on chromosome 7 and plays a
 CC role in pigmentation. A portion of the full-length LCR sequence,
 CC designated HS (ABV76779) was identified as a DNaseI hypersensitive region
 CC that was found to be deleted in the mutant tyrosinase allele chinchilla-
 CC mottled, which confers a variegated pigmentation phenotype. The full-
 CC length LCR, or its functional fragments or analogues may be incorporated
 CC into animal cell expression constructs to protect against the effects of
 CC transgene integration at different locations in the genome. The LCR-
 CC derived sequence reduces the variability of transgene expression and
 CC associated with integration at different chromosomal locations, and
 CC ensures optimal expression (as regards time, place and quantity),
 CC regardless of the site of integration. The present sequence represents
 CC the full-length LCR of the mouse tyrosinase gene. In addition to its
 CC protective effect against variable transgene expression, this sequence
 CC also contains an enhancer element (denoted as Box AB in the
 CC specification); however, this enhancer is not required for protection
 CC against transgene expression variability
 XX SQ Sequence 3711 BP; 1037 A; 595 C; 650 G; 1429 T; 0 U; 0 Other;
 Query March 11.0%; Score 249.899; DB 1; Length 3711;
 Best Local Similarity 62.2%; Pred No. 89;
 Matches 574; Conservative 0; Mismatches 296; Indels 53; Gaps 12;
 QY 292 TCTATTTCTGATTTCTATCTGCTCAATTTTAACTCAGTAGTCAGTGTGTTGCTTCC 351
 DB 617 TTCTTTCTTTATTTCTGCTGACCTCAAGTCATCCAGTAGGAGTGTGCGAGCTTC 676
 QY 352 ATAAGTTCTAGTATTTCTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 411
 DB 677 ATGAGTTCTAGTATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720
 QY 412 GGTGTCAGATAGGACATAGAGTATTTATTTCAATTTGTTTATCTGTCGAGACTTGGT 471
 DB 721 TGTGTCATTAAGACACAGGCAATTTTCAACTTTCTTATATCTGTTGAGGCTTATTT 780
 QY 472 TGTGTTGAATATGATTTCAATTTTGGAG-AGTTTCATAGGCTGCTGACAGAGATACA 530
 DB 781 TGTGAACAACTATATGATCAATTTTGGAGAAAGTTTAAAGGATTTGCTGAGAAAGGTGCC 840
 QY 531 GTCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 590
 DB 841 TCTTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 886
 QY 591 CAGTTAGTCTCAGCATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 650
 DB 887 CTTTACCTCCATATTTCTTAGTTTAG-TTTTTGTCTAGATGACATGCTATTTCAATGAG 945
 QY 651 AATGGGTTATCAAGTAGCCCACTATCTGTGTGAG-GTCAATATGATGATTTTAGTGTG 709

DB 946 AGTAGGATATTGAAGTCTCCCACTATTAAATGATTATTAGATTCAATTTGTGATTGAGCAT 1005
 QY 710 AGCTGTGCTGTTTATGAACATTGGGTGACATTTGTTTGGTGCATAGACATTAAAGAATT 769
 DB 1006 AGTAATGTTTCACTT-CCAAATGTTTACCTTGTCTTTGGGACATAGATGTTAAGNACT 1064
 QY 770 GCAATGTCCTCTTGG-TGGAATTTTCTTTGATGCCCTATGATATCTTCCCAATCTCAT 828
 DB 1065 GAAGTGTATCATGGCTAAATTTTCTTTGATGAGTATGAAATGCAATTTTCCCATCTCTA 1124
 QY 829 CTGCTTAGTTTGGGTTTAAAGCT-ATTAGTCAGATATTAAATGACTGTCATCGGCTTGC 887
 DB 1125 TTGACTTAATTTCTGTAAGTGTGATTTGCTAGATATTAGATAGCTAAACAGCTTGC 1184
 QY 888 TTCTTAGGCCAATTTGCTTTAGAAATCTTTTCCATCTTTTACTCTTAAGGTGATGCTAT 947
 DB 1185 TTCTTGAGTGCACTGCTTGGAAATATTTTTCACACCTTAACCTGGAGGATTTGCTAT 1244
 QY 948 C--CATGTTAGTGTCTTTTGGATGCAGCAGTAGGATGATCTTGTTTTCATATCCA 1005
 DB 1245 CTTGATGTTGAGTCTATTTTGTACTCAGTGAATGATGATCTCTTTTGCATCCA 1304
 QY 1006 TTCTGTATACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGTTGAGA 1065
 DB 1305 CTCTGTTTCCCTGTGTGTTT-----ATTGAGTAATTGAGTCCATTTACTTTGAGA 1355
 QY 1066 ATTATCAATGAGCAGTGTGTTGGATCTTGTATCTTTCACCTTGTGAGTGTGTGTG 1125
 DB 1356 AATATTAAATGATCAATGATTTCTTCAATTTCTGTTATTTTG-----GTGGGTGGTGTAG 1410
 QY 1126 TGTGCTGT 1184
 DB 1411 TAGTGATAGTTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1470
 QY 1185 CTGTGCTGT 1207
 DB 1471 CTGTATCATGATTCCTATATGC 1493

RESULT 40

AAC64370

ID AAC64370 standard; DNA; 125910 BP.

XX AC AAC64370;

XX AC AAC64370;

XX DT 07-FEB-2001 (first entry)

XX DE Human KCNQ5 (KCN6q) gene sequence SEQ ID NO:1.

XX KW Human; KCNQ5; KCNQ6; chromosome 6; voltage-gated potassium channel;

XX KW Stargardt-like macular dystrophy; cone-rod macular dystrophy;

XX KW Salla disease; ophthalmological; auditory; central nervous system;

XX KW cardioactive; anticonvulsant; gastrointestinal; muscular active;

XX KW age-related macular degeneration; macular degeneration; deafness;

XX KW epilepsy; neuropsychiatric disorder; heart disorder; muscle disorder;

XX KW gastrointestinal disorder; ds.

XX OS Homo sapiens.

XX OS WO200061606-A1.

XX PD 19-OCT-2000.

XX PF 10-APR-2000; 2000WO-US009587.

XX PR 14-APR-1999; 99US-0129274P.

XX PA (MERI) MERCK & CO INC.

XX PI Petrukhin K, Caskey CT, Li W, Metzker ML;

XX PI WPI; 2000-647417/62.

XX DR P-PSDB; AAB24241.

XX Voltage-gated potassium channel KCNQ5 DNA and protein, for identifying
PT inhibitors and activators which can treat e.g. Stargardt-like macular
PT dystrophy, cone-rod dystrophy, Salla disease, deafness, and epilepsy.
XX Claim 3; Fig 1; 99pp; English.
XX
CC The present sequence represents the human KCNQ5 (also called KCN6q) gene,
CC which encodes a voltage-gated potassium channel protein. Human KCNQ5 has
CC ophthalmological, auditory, central nervous system (CNS), cardioactive,
CC anticonvulsant, gastrointestinal and muscular active activities.
CC Sequences and methods from the present invention are useful for
CC identifying activators or inhibitors of KCNQ5 protein. These activators
CC and inhibitors are useful for treating Stargardt-like macular dystrophy,
CC cone-rod dystrophy, Salla disease, age-related macular degeneration,
CC other forms of macular degeneration, deafness, epilepsy, and different
CC forms of neuropsychiatric, heart, gastrointestinal, and muscle disorders.
CC Stargardt-like macular dystrophy and cone-rod dystrophies are located at
CC chromosome 6q
XX
SQ Sequence 125910 BP; 40132 A; 24180 C; 23166 G; 38360 T; 0 U; 72 Other;
Query Match 11.0%; Score 249.898; DB 1; Length 125910;
Best Local Similarity 51.0%; Pred. No. 15;
Matches 878; Conservative 0; Mismatches 766; Indels 77; Gaps 13;
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

RESULT 41
AAS63340
ID AAS63340 standard; DNA; 6207 BP.

1013 ACCAGTATCTTTCTAGAGAAATTAAGATCAATGAGTCATTGATGTTGAGATTTATCA 1072
66128 ATTCTATATCTTTT-----AAGTGGAGCATTTAAGCCATTACATTCACGTTAGTA 66179
1073 ATGAGCAGTGTGTTGGGATCTTGTATCTTCGACTGTGGAAGTGTGTGTGTGTGT 1132
66180 TTGAGATGTGAGTACCATTCATTCATATGC-----TATTTGTTCCCTGTATACCT 66232
1133 GT 1192
66233 TGGTTTTTTTGTGTTTTGTGTTTTTAAATGTAATTTTGTGTTTATAGTCTCTGTGA 66292
1193 TGTGTGTGTGTGTCTCTCCCTCTTTGAAATTTTGGCTGGAATTTATTTATTTATCA 1252
66293 GATTTATGCTTTTAAAGAGGTTCTGTTTGTGTTGAGTGTTCAGGATTTGTTTCAAGAT 66348
1253 TTTTCTGTAATGTGGGTAAATCTTTAGATCAATCTTTAGTCAAGTTTTCCTTAGCTTC 1312
66349 -----TTAGAGTCTCTTTTAGCAGTTCTGTAGTGG 66379
1313 TGCATTTGAAGATAGATATTTCTTTACATCTGATTTTATCTTAGAATGCTTTCTTTCTCC 1372
66380 TGGCTTGGTAGTGTGAATTTCTCTCAGAAATTTGTCTGTGAAAGAGCTGATCTTTTCC 66439
1373 AACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCTGCACATCTGTAGTCTCT 1432
66440 TTCTATATGAAGCATAGTTTGTGTTGATACAAAATTTCTGGCTGGTAAATCTTTTGTCT 66499
1433 TGGAGTCTGTAGCATCTGTGAGGCTCTTACATTTTGTGATTTTCTTTTGGAAAGT 1492
66500 GAGGAGGCTGAAGACAGGCCCAATACCTTCTAGTATAGGTTTCTGCT-GAGAAAT 66558
1493 CAGGTGTAATTTCTATACATCTGCCTTTATATGTTAAATGTTGCTTTTTCCTGCACT 1552
66559 CTGCTGTTAACTAATAGGTTTATTTATAGGTTTACAGGTAATTTTGT-CTCACTGT 66617
1553 TTTAATATTTCTTTCTGTTCTATCTTTAGTGAATTTGATTTATGCACTGTGCGGAG 1612
66618 CTAAATTTCTTTCTTTGCTTAACTTTAGATTAACCTGATGCAATGTCCTAGTGTAT 66677
1613 TTTCTTTCCGGTCCAATCTATTGTTGTTTGTATGTTTGTACCTTTGATAGGCAATC 1672
66678 GATCTTTTGTCAATGAATTTCCAGGTTTCTTTGCTGCTTTGTTATTTGGATGCTAGG 66737
1673 TCTTTCTCAGGTTAGGAAATTTTCTTTTGTGTTTCTTTGAAATATTTTCCCTGCTT 1732
66738 TCTTAGCAATACTGGGAAAGTTTCTCTGTTATTTCCCTGAAATATGTTTCCAACTT 66797
1733 TTGACCTGCTTTCTTCCCT-----TCCCTCTATTCTTGTGTTTGTGCTAGTG 1781
66798 TTAGATTTCTCTTCTTCTCAAGACATGATATTTATAGCTTTGGTCACTAACATAA 66857
1782 TCTCTGGCTTCTGATGTTTATGCTGGATTTATTTAGACTTAACATTTTCTTTTGACC 1841
66858 TCCAGACTTCTGGAGGCTGTGTTTCTGATTTTCTTATTTCTTTTCTTTTCTTTTCT 66916
1842 AAGGATCATTTCTTCTATCTTCTGCTGCTGCTGAGATTTCTCTCTCTATCTCTTGT 1901
66917 GAGTGGTTAAGTGAAGACCTTGTCTTAGACTCTGAATCTTTCTTCTTCTTCTTCTTCT 66976
1902 ATTCTGTGAGGCTTGTCTCTGAGTTCTGTGGGTTCTGTGGGTTCTTAATTTTCTTCTCA 1961
66977 ATTCTATTGCTGAGACT--TTCAGAGCATTTTGCATTTCTATAGGGTGTCCAAAGTTT 67034
1962 GATTTCTCTCAGTTGGGTTTGTGTTTATTAATCTTATTTCC 2002
67035 ACTGAAGTTTGTAGCTGTTTATTTATTTAGCTATCTATTTCC 67075


```
QY 586 GACATCAGTTAGCTCCAGCATTTCTCTGTTTCGTTTTTTTGTGTGAGTACCTAAGTCTTG 645
DB 21320 GGTATAGTTTAAAGTCCATTTTTCCTTTG-TCGACTTTCTGCTTGAAGACCTGCTAGTG 21378
QY 646 GAGAGAAATGGGGTATTGAAGTAGCCACTATCTGTGTG-TGAGGTCAATATGTGATTTTA 704
DB 21379 CTGCCAGTGGAGTATGAAGTTCGCACTATTTTATGTTGCCATCTAGCTCACTTCTTA 21438
QY 705 GCTGTAGCTGTCTGTTTATGAACCTTGGGTGACATGTGTTGGTGTGCATAGACATAA 764
DB 21439 GGTGTAGTAGTGAAGTGTATATAAATTTGTGAGCTCCAGTCTTAGGTAAATATATATTA 21498
QY 765 GAATTTGCAATGCTCTCTGGTGA-----TTTTCTTTGATGCTATGTAGTATTC 815
DB 21499 GGAATGTGACATTTTCTGTGGACTAATCTTTTAAATTTATATATGACCCACTTTGT 21558
QY 816 TTCCCAATCTCATCTGCTTAGTTTGGGTATTAAGTCTATTAAGTCAAGATATTAATAAGT 875
DB 21559 CTTTCTTTTAACTGCTGTGCTCTGAAGTCTGTTTGTAGTCTGATATAAGACTAGC 21618
QY 876 GTATCGGCTGCTCTTAGGCCCATTTGCTTAGAATATCTTTTCCATCCCTTTTACTCTAA 935
DB 21619 TACTCTGCTGCTTGGTTTCCATTTGCGTGGAAATATTTTCCACCCCTTTACTT-TAA 21677
QY 936 GGTGATGTCTATCCATG---GTAGTGTCTTTTTTGGATGACAGTGTAGGATGATCTT 992
DB 21678 GTTTATATGAATCTCATGTGTCAGGTGAGTCTCTTGAAGACAGATACACTTGGTTGGT 21737
QY 993 GTTTTCATATCATCTGTTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTC 1052
DB 21738 GGAATTTTATCCATCTGCCATCTGCACTCTTCTAAGTGGAGCATTTAGA-----TC 21789
QY 1053 ATTGATGTTGAGAAATTAATCAATGACAGTGTGTTGGATTCCTGTTATCTTGACATTTGTG 1112
DB 21790 ATTACGTTTATGTTAGTATGAGATTTGAGTACTGTTCAATTAATCATGCTA----- 21844
QY 1113 AAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1172
DB 21845 -----GTTGCTTC 21852
QY 1173 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1232
DB 21853 CTTAAATCTTTTCCCATTTGTTTATGTTTATAGGCTTGTGTGTGTGTGTGTGTGTGTGTGT 21912
QY 1233 GGAATTTATTTATTCATATTTCTTGAATGTGGTAACATCTTTTGAATTTGAATTTT 1292
DB 21913 AGGAGTCTATTTTGGTGTATTT-----TGAGTCTTCTCCAAATTTGGAACCTCT 21966
QY 1293 CTCCTAGCTCTTTAGTCTGCAATTTGAAGATAGATATCTTTTACATCTGATTTATCT 1352
DB 21967 CTTACTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 22026
QY 1353 TAGAATG--TCTTTCTTCTCCAACTATTGTGACAGAAAGTTTCTCT---AAGTGCAGTA 1407
DB 22027 GAAAGGAATTTTAACTCTCTTCAATTTATGAAGCTTAGTTTGTCTGATATAAAAAA 22086
QY 1408 GTCTGCTGTGACATCTGTAGTCTCTTGTGAGTCTGTAGCAGATCTGTGACAGGCCCTTCTTA 1467
DB 22087 TTTGTGGTGTGACAAATTAATTTGTTTTCAGGAGGCTAAAGATGGGACCCCAATACATCTGC 22146
QY 1468 CATTTTGTAGTTTCTATTGGAAGAGTCAGGTGTAAATCTTAATACATCTGCCCTTTATATGTT 1527
DB 22147 CTTGTAGGTTTCTGCTTAGA-----AATAGGTTTCTTTATAGATT 22189
QY 1528 AATTGGTCTTTTCCCTTGCATCTTTTAAATCTTTTGTGTTGTTATACATTTTAGTGA 1587
DB 22190 AACTGATGCTTTGTTTCCATAGCTCTTAAGATTTTCTCTGTCTGTCTGACTTTAGACAA 22249
QY 1588 TTTGATTTATGCACTGTGGGAGTTCTTTTCCGTCCTCAATCTATTTGTGTGTTTT-GT 1646
DB 22250 CTTGATCTGATGCTGTACATGATAATCTTTTGTGATGAATTTCCAGATGTTTTCTT 22309
QY 1647 ATGCTTCTTGTACCTTGTAGTGGCATCTCTTTCTCAAGGTTAGGAAATTTTCTTTTGTG 1706
```

```
DB 22310 AAGCTCTTGTATTGAATGCTAGATCTCTAGCAAGCTGGGAGCTTTCCAGTT--- 22366
QY 1707 TTTTCTTGAAAAATATTTTCCCTGCTTTTGGACCTGCTTCTTCCCTTCTCTATTCCTTT 1766
DB 22367 GTTTCCTCAATAAGTTTTCCAAATTTTAGATTTCTCTTTATTTCTGGGAATACCAATT 22426
QY 1767 GGTTTTGCATAGTCTCTGCTGCTGCTTCCCTGGATGTTTATGCTGCTGATTTT----- 1819
DB 22427 ATTTTGTAGTTTGGCCATTTAAATTTCAATTTCTTGTGCTTTTGTGCTTTTGTTCATTAC 22486
QY 1820 ----AGACTTAACATTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTGCTTCACTGC 1875
DB 22487 TCAATTTTAAATTTGCTTGTCTGATTTGGTTAAATCAAAACCTTGTCTTCGAGCC 22546
QY 1876 CTGAGATCTCTCTTCTATCTCTGTTATCTGCTGAGTGAGGCTTCTCTGAGGTTCTCTG 1935
DB 22547 TTGAAGTTTCTTCTTCTTCTGTTTGTAGTCTATGCTGACACTTCTAGTGCATTTGTG 22606
QY 1936 TTGGGTTCTTAATTTTCTTCAATTTCCAGA-----TTTCTTCAAGTTTGGTTTGTATTA 1991
DB 22607 TTTCTTAGCGTATTTTCAATTTCCAAAATATGATTTGTTCTGCGTTGTGATCTG 22666
QY 1992 ATTCATTTCCACTTTCAGGCTCTGAAATGTTTACTCATTTTCTCTCCAGTATTTACAT 2051
DB 22667 TTTCTCTGGATAATTTTTCATCCATATCTGTATTTGTTTAAATTCATTCAGTTGTT 22726
QY 2052 TTTTCATAGGTTTCTTTAAT 2070
DB 22727 TTTACCTTTCTCTGGTAT 22745

RESULT 43
AAC64370/c
ID AAC64370 standard; DNA; 125910 BP.
XX AC AAC64370;
XX DT 07-FEB-2001 (first entry)
XX DE Human KCNQ5 (KCN6q) gene sequence SEQ ID NO:1.
XX KW Human; KCNQ5; KCNQ5q; chromosome 6; voltage-gated potassium channel;
Stargardt-like macular dystrophy; cone-rod macular dystrophy;
Salla disease; ophthalmological; auditory; central nervous system;
cardioactive; anticonvulsant; gastrointestinal; muscular active;
age-related macular degeneration; macular degeneration; deafness;
epilepsy; neuropsychiatric disorder; heart disorder; muscle disorder;
gastrointestinal disorder; ds.
XX OS Homo sapiens.
XX FN WC200061606-A1.
XX PD 19-OCT-2000.
XX PF 10-APR-2000; 2000MO-US009587.
XX PR 14-APR-1999; 99US-0129274P.
XX PA (MERI ) MERCK & CO INC.
XX PI Petrukhin K, Caskey CT, Li W, Metzker ML;
XX DR WPI; 2000-647417/62.
XX P-PSDB; AAB24241.
XX PT Voltage-gated potassium channel KCNQ5 DNA and protein, for identifying
inhibitors and activators which can treat e.g. Stargardt-like macular
dystrophy, cone-rod dystrophy, Salla disease, deafness, and epilepsy.
XX PS Claim 3; Fig 1; 99pp; English.
XX XX
```

The present sequence represents the human KCNQ5 (also called KCNQ6) gene, which encodes a voltage-gated potassium channel protein. Human KCNQ5 has CC ophthalmological, auditory, central nervous system (CNS), cardioactive, CC anticonvulsant, gastrointestinal and muscular active activities. CC Sequences and methods from the present invention are useful for CC identifying activators or inhibitors of KCNQ5 protein. These activators CC and inhibitors are useful for treating Stargardt-like macular dystrophy, CC cone-rod dystrophy, salla disease, age-related macular degeneration, CC other forms of macular degeneration, deafness, epilepsy, and different CC forms of neuropsychiatric, heart, gastrointestinal, and muscle disorders. CC Stargardt-like macular dystrophy and cone-rod dystrophies are located at CC chromosome 6q

SQ Sequence 125910 BP; 40132 A; 24180 C; 23166 G; 38360 T; 0 U; 72 Other;

Query Match.	6.2%;	Score 141.4;	DB 1;	Length 125910;
Best Local Similarity	52.9%;	Pred. No. 25;		
Matches 531;	Conservative 0;	Mismatches 431;	Indels 42;	Gaps 11;
QY	292	TCTATTCTTGATTTCTATCTCTGGCFCATTTTAACTCAGTAGTCAGTGTGTTGGTTTCC	351	
Db	71917	TTTGCTCTTAATCTCTTCATTGACCCATGGTCATTGAGGAGCATATGTTTAATTATC	71858	
QY	352	ATAAGTTTGAAGTTTCTGTTGTTTCTGTTGTTGTTGTTTATCTATAGATTTAAGCGT	411	
Db	71857	ATGTGTTGTGATAGTTTCCAAATTTTCTCTGTACTGATTTT---TASTTTTCATTCAA	71801	
QY	412	GGTGGTCAGATAGACATAGAGTATTATTTCAATTGTCTTTTATCTGTGCGAGACTTGCTT	471	
Db	71800	TGTGGTCAGAGAAGATACTTGATATTAATTTCAAATTTTC-TGAACTGCTTAAGACATGTTT	71742	
QY	472	TGTTTTGAATATGTAATTCAAAT-----TTTGAGAGTTTCATAGGGTCTGAC	519	
Db	71741	TGTGGCTTAACATGTGGCTATCTCTGATCTATCTTGAATAAATCCATGTGTTAGGAG	71682	
QY	520	AAGAAGTACAGTCCT---TGTTGTTGGTGAATAGTCTGTAATATCTCTAGGTCCA	575	
Db	71681	AAGTGTGGAGTATTTCTGCAGCTGTTGAGTGAATGTTCTATAAATTAATTTATAGTCC	71622	
QY	576	CTTGGTTTATGACATCAGTTTAGTCCAGCATTTCTCTGTTTGGTTTTGTTTGAGATGAC	635	
Db	71621	ATTTGGTCTATTGTGCAGATTAAGTTTGTATGTTTCTTTGTTGATTTCTGTCTGGATGAT	71562	
QY	636	CTAACTGTTGGAGAGAAATCGGGTATTGAAGTAGCCACATCTCTGTG-TGTCAGGTCAATA	694	
Db	71561	CTGTCCAAATGCTGAAAGTTGGGTGTTGACGTTTCAGCTATTAATTTGACTCTGGGGTATGTC	71502	
QY	695	TGTGATTTTAGCTGTAGCTGTCTGTGTTTTATGAACCTTGGGTGACATTTGTTTGGTGCA	754	
Db	71501	TCCTCTTTTGTCTTAATAATATTTGCTTTATATATCTGGGTGCTCCAGTGTGGTGCA	71442	
QY	755	TAGACATTAAGAAATTCGAATCTCTCTGGTGGATT-TTCCCTTGATGCCCTATGTAGTAT	813	
Db	71441	TATATATTTATAGTTGTATATCTCTGCTGAATGATTCCTTTATCATTTATATAATGA	71382	
QY	814	TCTTCCCAATCTCATCTGCTTAGTTTGGTTTTAAAGTCTATT-AGTCAGATATTAAAATG	872	
Db	71381	CCTTCTCTGTCTCTTTTGATAGCTTTTGCTTGAAACCTATTCTGCTGATCTAAGTGA	71322	
QY	873	ACTGTATCGCTGCTCTTTAGGGCCATTTGCTTAGAATATCTTTTC-CATCCCTTTACT	931	
Db	71321	GCTACTCTGCTCTTTTCGGTTTCCATTTGATGGGATATCTTTTTTATTCCTCTTATT	71262	
QY	932	CTAAGGTGATGCTATPCCAT---GGTAGTGTCTTTTTTGATGCAGCAGTAGGATGGA	988	
Db	71261	TTCAATCTATGTAGTCTTTATAGATGAAGTATGTTTCTTGTAGTCAACAGATCATTTGG	71202	
QY	989	TCCTGTTTTCATATCAATCTCTGTACCAGTATCTT-----TTTCTAGAGAAAT	1037	
Db	71201	TCCTGCTTTTAAATTTAAATTTTAATCAATTCGAATCAATTCATCTCTTTTGAT	71142	
QY	1038	TAAGATCATTTAGTCAATGATGTTGAGAAATTAATCAATGACGAGTGTTTGTGGATCTTGT	1097	

[illegible]

[illegible][illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 07:11:15 ; Search time 3651 Seconds
(without alignments)
3.978 Million cell updates/sec

Title: US-10-664-775-4

Perfect score: 2279
Sequence: 1 gatcactcctctagtgaaag.....ttgtaattctagtgctgat 2279

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 39 seqs, 3186700 residues

Total number of hits satisfying chosen parameters: 78

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : rni4.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	588.3	25.8	29604	1	US-08-781-891-207
2	588.3	25.8	29604	1	Sequence 207, App
C	3 565.399	24.8	51259	1	US-08-781-891-209
C	4 565.399	24.8	51259	1	Sequence 209, App
5	521.797	22.9	51259	1	US-08-781-891-209
6	521.797	22.9	51259	1	Sequence 209, App
7	323.899	14.2	16442	1	US-08-781-891-208
8	323.899	14.2	16442	1	Sequence 208, App
9	253.097	11.1	193303	1	US-09-497-855A-37
10	253.097	11.1	193303	1	Sequence 37, Appl
C	11 252.597	11.1	392000	1	US-09-497-855A-44
12	251.899	11.1	10409	1	US-10-027-983-11
13	248.997	10.9	168575	1	US-08-772-440-33
14	242.997	10.7	116592	1	US-09-426-230-1
15	239.897	10.5	786431	1	US-09-818-512-3
16	239.797	10.5	32042	1	US-09-751-389-3
17	239.797	10.5	32042	1	Sequence 44, Appl
18	208.398	9.1	202001	1	US-09-340-620A-63
C	20 204.799	9.0	99916	1	US-08-742-185-102
C	21 200.5	8.8	148567	1	US-09-734-674-3
C	22 200.5	8.8	148567	1	Sequence 3, Appl
C	23 191.799	8.4	168575	1	US-09-801-876B-3
C	24 186.899	8.2	2116	1	US-10-254-869-3
C	25 186.899	8.2	4646	1	US-09-426-230-1
26	186.1	8.2	392000	1	US-09-377-437-11
27	177.9	7.8	99916	1	US-09-377-437-67
C	28 176	7.4	786431	1	US-10-027-983-11
C	29 169	7.4	161652	1	US-09-816-095-3
C	30 149.3	6.6	70000	1	US-09-751-389-3
C	31 145.4	6.4	64467	1	US-09-497-855A-40
32	142.8	6.3	1540	1	US-09-851-896-3
C	33 136.8	6.0	9573	1	US-08-803-671B-3
					US-09-220-132-168

34	125.5	5.5	50000	1	US-09-146-053-4	Sequence 4, Appli
C	35 117.1	5.1	3033	1	US-09-525-160B-9	Sequence 9, Appli
C	36 117.1	5.1	3213	1	US-09-525-160B-4	Sequence 4, Appli
C	37 116	5.1	116592	1	US-09-818-512-3	Sequence 3, Appli
38	115.3	5.1	38564	1	US-09-734-673-3	Sequence 14, Appl
39	108.3	4.8	8224	1	US-09-010-398-14	Sequence 14, Appl
40	108.3	4.8	8224	1	US-09-366-260-14	Sequence 14, Appl
41	106.2	4.7	468	1	US-09-702-705-604	Sequence 604, App
42	106.2	4.7	468	1	US-09-736-457-604	Sequence 604, App
43	106.2	4.7	468	1	US-09-614-124B-604	Sequence 604, App
44	106.2	4.7	468	1	US-09-671-323-604	Sequence 604, App
45	106.2	4.7	468	1	US-09-589-184-604	Sequence 604, App

ALIGNMENTS

RESULT 1
US-08-781-891-207
; Sequence 207, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 207:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29604 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-207

Query Match	25.8%;	Score 588.3;	DB 1;	Length 29604;
Best Local Similarity	62.9%;	Pred. No. 0.24;		
Matches 1170;	Conservative	0;	Mismatches 567;	Indels 123; Gaps 18;
Qy	292	TCATTTCCTGATTTCTATCTTGCGCTCATTTTAACTCAGTAGTGTGTTGGTTTCC	351	
Db	23722	TTTCTTTCTTTATTCCTCTCTTGACCAAGGTATCATTCAGAAGAGTGTATTCACTTCC	23781	
Qy	352	ATAAGTTGTGAAGTTTCTCTGTTTCTGTGTTGTTGTTATCTAGATTTAAGCTGT	411	
Db	23782	ACGTGAATCTGCGCTTTCATTATTATGTTGTTATTTGAAGATC---AGCCTTAGGCCAT	23838	

ATTORNEY/AGENT INFORMATION:
 NAME: Mcmasters, David D.
 REGISTRATION NUMBER: 33,963
 REFERENCE/DOCKET NUMBER: 240052.419C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 207:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 29604 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 207:
 US-09-618-166-207

Query Match 25.8%; Score 588.3; DB 1; Length 29604;

Best Local Similarity 62.9%; Pred. No. 0.24;

Matches 1170; Conservative 0; Mismatches 567; Indels 123; Gaps 18;

Qy	292	TCTATTCCTGATTTCTATCTGGCTCATTTTAACTCAGTAGTGAGTTGTTGGTTCC	351
Db	23722	TTTCTTTCTTATTCCTTCCCTTGACCAAGGTATCATTGAGAAGAGTGTTATTCAGTTCC	23781
Qy	352	ATAAGTTTGTAAAGTTTCTGTGTGTTCTCTGTTGTTGTTGTTTATCTAGATTAAAGCTGT	411
Db	23782	ACGTGAATGTTGGCTTTCCATTAATTATGTTGTTATTGAAGATC--AGCCTTAGGCCAT	23838
Qy	412	GGTGTGATAGACATAGAGTATTTTCAATGTCTTTTATCTGTCTGAGACTTGCCTT	471
Db	23839	GGTGTGTGATAGATACATGGGCAATTTCAATATTTTGTATCTATTGAGGCTCTGTT	23898
Qy	472	TGTTTGAATATGTAATCAATTTGGAGA--GTTTCATAGGGTGGCTGACAGAAGGTAC	529
Db	23899	TGTACCAATTTATGTTGTCATTTTGGAGAAGTCCCGTGGGTGCTGAGAAGAGTAT	23958
Qy	530	AGTC-TTGTGTTTGGTGAATAGTCTGTAATATCTCT-AGCTCCACTGGTTTATGA	587
Db	23959	ATCCTTTTGTGTTAGGATAAAATGTTCTGTAGATATCTGTCAAGTCCATTTGTTTCATA	24018
Qy	588	CATCAGTTAGCTCCAGCATTTCTGTCTCGTTCTGTTTGTGAGATGACCTAACTGTGGA	647
Db	24019	CTTCTGTGTAGTTTCACTGTGCTCCCTGTTTAG-TTTCGTGTTCCACGATCTGCTCTTGA	24077
Qy	648	GAGAAATGGGTATTTGAAGTAGCCCACTATCTGTGTGTGAGGT-CAATATGTGATTTTACG	706
Db	24078	GAAAGTGTGTGTTGAAGTCTCCACATTTATTGTGTGAGGTGCAATGTATGCTTTGAGC	24137
Qy	707	TGTAGCTGTGCTGTTTATGAACCTTGGTGACATTTGTTGTTGTCATGACATTAAGA	766
Db	24138	TTTACTAAAGTGTCTCTAATGAATGTGGCTGCCCTTGCAATTTGGTGGGTAGATATTCAGA	24197
Qy	767	ATTGCAATGTCTCTTGGTGGATTTT-CCTTTGATGCTATGTAGTATTTCTTCCCAATCT	825
Db	24198	ATTGAGTGTCTCTTGGAGATTTTACCTTTGATGATGATGAAGTGTCCCTCTCTGTCT	24257
Qy	826	CATCTGCTTAGTTTGGGTTTAACTCTA-TTATGTCAGATATAAAATGACTGTATCGGCT	884
Db	24258	TTTTTGATAAATTTGGGTGGAAGTCGATTTTATCCGATATAAAATGGCTACTCCAGCT	24317
Qy	885	TGCTTCTTAGGCCATTTGCTTAGATA-TCTTTTCCATCCTTTTACTCTAAGGTGATGT	943
Db	24318	TGTTTCTTCACTCCATTTGCTTGGAAATTTGTTTCCAGCCCTTTTACTCTGAGGTAGTGT	24377
Qy	944	CTATCCAT--GGTAGGTGTCTTTTGGATGATGAGGTAGGATGGATCTTGTTTTCAT	1000
Db	24378	CTGTCTTTTCCCTGAGATGGTTTCTCTGTAAGCAGCAGATGTTGGTCTCTGTTGTGT	24437
Qy	1001	ATCCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCATTGATGT	1060
Db	24438	AGCCAGTCTGTATGTCTATGTCTTTTATTGGGGAATTGAG-----TCCATTGATAT	24489
Qy	1061	TGAGAAATATCAATGAGCAGGTGTTTGTGGATTCTTGTTATCTTGCACTTGTGGAAGTGTGT	1120

RESULT 3

Db	24490	TAAGAGATATTAGGAAAAAGTAATTTGCTTCCCTTTTATTTTGTGTAGATTGGCA	24549
Qy	1121	GT	1180
Db	24550	TTCTGTTC-TGTGGCTTCTCTTTTGGTTG-----	24582
Qy	1181	GT	1240
Db	24583	-----TTGAATGAT	24591
Qy	1241	TTATTTATCATATTTCTTGAATGTGGGTAAACATCTTTAGATTGAAGTTTTTCTCTCAGC	1300
Db	24592	TACTTCTTGGTGTCTTAGGGCGTATTTCCGTCTTGTATGCTTCTTTCTGTATT	24651
Qy	1301	CTTCTTT--AGGTCTGCATTTGAAGATAGATATTTCTTTACATCTGATTTATCTTGAAT	1358
Db	24652	ATCCTTTGAAGGGCTGGATTCGTGGAAGATATTTGTGAAATTTGTTTGTCTGTGGAAT	24711
Qy	1359	GTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCTCA	1418
Db	24712	ACTTGGTTCCTCATCTATGTAATTGAGAGTTTGGCCTGGTATAGTAGCCTGGGCTGG	24771
Qy	1419	CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGGGCCTTCTTACATTTTGAGTT	1478
Db	24772	CATTTCTGTTCTTCTAGTTTCTGTATAACATCTGTCCAGGCTCTTCTGGCTTTCATAGTC	24831
Qy	1479	TCTATTGGAAGTCAAGTGTAAATCTTAATACATCTGCCCTTATATGTTAAATGGTCTTT	1538
Db	24832	TCTGGT-GAAAAAGTCTGGTGTAAATCTGTATAGGCCCTTCTTTATATGTACTT-GACCTT	24889
Qy	1539	TTTCCCTTGCATCTTTTAAATATTTCTTTCTTGTGTTCTATCTTTTAGTGATTTGATTATTA	1598
Db	24890	TCTCCCTTACTGCTTTTAAATATTTCTATCTTTTAGTGCATTTGTTGTCTGATTAATTA	24949
Qy	1599	TGCATCTGGGAGTTTCTTTTCOGGTCCAAATCTATTTGGTGTGTTGTTGATGTTCTGTGA	1658
Db	24950	TGTGTCGGGAGAAATTTCTTTCTGGTCCAGTCTATTTGAGGTCTGTAGGCTTCTTTGA	25009
Qy	1659	CCTTGATAGCATCTCTTTCTCAAGGTAGGAAATTTTCTTTTGGTTTCTTTGAAAAA	1718
Db	25010	TGATCAAGGCACTCTCTTTTATGTTTGGGAAGTTTCTTCTATATTTTGTGGAAGA	25069
Qy	1719	TATTTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCTCTATCTCT	1764
Db	25070	TATTAGCTGGCCCTTTTAAAGTTGAAAAATCTTCAATCTCATCAATTCCTATTAATCCGTAGCT	25129
Qy	1765	TTGGTTTTTGCATAGTGTCTCTGGCTTCTGATGTTTATGCCCTGGATATTTTAGACT	1824
Db	25130	TTGGTCTTCTCATTTGTGCTCCTGGATTACCTGATGTTTGGATTTAGGATCCTTTTGCATT	25189
Qy	1825	TAAACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGCAGATTC	1884
Db	25190	TTGATATTTCTTTGACTGTTGTGCGATGTTCTCTATGGAATCTTCTGCACCTGAGATTC	25249
Qy	1885	TCTTCTTATCTCTTGTATCTGTGTCAGTGGCTGTCTCTGAGGTTCCTG-TTGGGTTTC	1943
Db	25250	TCTTCTTCCATTTCTTGTATCTGTTGCTGATCTCGCATCTATGTTCCAGATCTCTTTC	25309
Qy	1944	TTAATTTTCTTCTTCCAGATTTCTTCAAGTTTGGTTTGTGTTTATTAATCTTATTTCCA	2003
Db	25310	CTAGGATTTCTATCTCCAGCTTGGCTCGCTTTGGTTTTTCTTATTTGTGCTACTTCCC	25369
Qy	2004	CTTTCAGGTCCTGAAATGTTTACTCATTTTCTCTCCAGTATT---TACATTTTCTAGG	2060
Db	25370	CTTTTAGTCTAGTATGGTTTTGTTTCATTTCCATCCACTGTTGGATGTTTTCTCTGT	25429
Qy	2061	TTTCTTTAATGATTTATTCATTTCTCTCTCAAGACCTTTTATGAAATTCATAAATGTA	2120
Db	25430	TTTCTTTAATGATTTTCTACCTCTTTGGCTGTGTTTTCTCTGCTTTTCTTTAAGGCTGTA	25489

1764 QY -----TTGGTTTTGATAGTGTCTCTGGCTTCTCGATGTTTTTATGCTGGAAT 1814
1815 Db ATCCTTAGGTTTCATTTCTCATTTGTGCTTATATTTCTCGATGTTTTGGTTAGGAGC 8940
1815 QY ATTTTAGACTTAAATTTTCTTGACCAAGGATCCATTTCTTCTATCTTCTCTACTG 1874
8939 Db TTTTTCGATTTGATTTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 8880
1875 QY CTGAGATCTCTCTCTATCTTCTGATTTCTGCTGAGGCTTCTCTGAGGTTTCT 1934
8879 Db CTGAGATCTCTCTCTATCTTCTGATTTCTGCTGAGGCTTCTGATATATGACTGCT 8820
1935 QY G-TTGGGTTCTAAATTTTCTTCTGATTTCTGCTGAGGCTTCTGAGGTTTCT 1993
8819 Db GATCTTCTCTAGGTTTCTGCTGCTGAGGCTTCTGCTGAGGCTTCTGATTTCTT 8760
1994 QY TCTATTTCCACTTTCCAGTCTCTGAAATGTTTCTTCTCAATTTCTCTCC---AGTATTACA 2050
8759 Db TCTAGTTCCATTTTAGATCTCTGATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8700
2051 QY TTTTCATAGGTTTCTTCTTCTGATTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 2109
8699 Db TCTCTCTAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8641

RESULT 4

US-09-618-166-209/c
; Sequence 209, Application US/09618166
; Patent No. 6583112

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui
Yu, Chang-En
Oshima, Junko
Mulligan, John T.
Schellenberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
WERNER'S SYNDROME

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed Intellectual Property Law Group

STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/618,166

FILING DATE: 17-Jul-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Mcmasters, David D.

REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 240052.419C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 692-6031

INFORMATION FOR SEQ ID NO: 209:

SEQUENCE CHARACTERISTICS:

LENGTH: 51259 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 209:

US-09-618-166-209

Query Match

Best Local Similarity

24.8%; Score 565.399; DB 1; Length 51259;

62.1%; Pred. No. 0.18;

Matches 1155; Conservative 0; Mismatches 576; Indels 128; Gaps 18;
292 QY TCTATTTCTTGAATTTCTATCTTGGCTTCATTTTAACTCAGTAGTGTGTTGTTGTTTCC 351
10412 Db TCTTCAFTTTCATTTCTGCTTTCGACCAAGGATCATTTAGTAGAGCGCTTTCAGCTTC 10353
352 QY ATAGTTTGTAAAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 411
10352 Db ATATGATGTGTGCTTTCCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 10293
412 QY GTGGGTTCAGATAGACATAGATATTTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 471
10292 Db GTCTCATGATAGGTTGATGAGGATTTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 10233
472 QY TGTGTTGAAATATGATATTTTCAATTTTGGAGA--GTTTCATAGGTTGCTGACAAAGGTAC 529
10232 Db TGTGACCAGCATATGCTGATTTTGGAGAAGTTTCCATGAGGTTGCTGAGAAGAAGGTAT 10173
530 QY AGTCTTTGTTGTTTGGTGAATAGTCTGTAATAATCT-CTAGTCCACATTTGTTTATGAC 588
10172 Db ATTTATTTGCTTTTGGATGACATGTTCTATATAATCTGTTAGATCCATTTGTTTCTATAAC 10113
589 QY ATCAGTTAGTCTCAGCATTTCTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTG 648
10112 Db ATCTGTTAGTTTCACTGTCTCTGATTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTT 10053
649 QY AGAATGGGTTATGAAAGTACCCACTATCTCTGCTGAGGT-CAATATGTTGATTTAGCT 707
10052 Db AGATGGGTTGCTGAAATCTCCACATTTATTTGATCGGTTGATGTTGTTGTTGTTGTTGTTG 9993
708 QY GTAGCTGTGCTTGTGTTTATGAACTTGGGTGACATTTGTTTGGTGCATAGACATTAAGAA 767
9992 Db TTAGTAAAGTTTCTTTTATGAAATGTTGGTCCCTTGCATTTGTAGCATACATGTTTCAGAA 9933
768 QY TTGCAATGCTCTTGGTGGAT-TTTCCTTTGATGCTATGATGATTTCTTCCAAATCTC 826
9932 Db TTGAGAGTTTCACTTGGCAGATTTTCTTTCACCATGATGAAGTCTCCTTCTTCTTCTTCT 9873
827 QY ATCTGCTTGTGTTTGGGTTT-----AAGTCTATTAGTCAGATATTTAAATGACTGTATC 880
9872 Db TTTTCTTTTGTGATTAATTTGTTGAGATTTGATTTATTTTCCATATTAGAAAGGCTACTCC 9813
881 QY GGCTTCTCTTTAGG-----GCCATTTGCTTGAATA-TCTTTTCCATCCTTTTACTCTA 934
9812 Db AGCTTGTCTTCTGGGAAACACACATTTGCTTGGAAAATTTGTTTCCACACTTGAACCTG 9753
935 QY AGTGATGCTATCCATG---GTAGTTCTCTTTTGGATGACGACAGTAGGATGATCT 991
9752 Db AGTAGTGTCTGCTTTTGTACCTGAGGTGCAATTTCTGTAAGCAAAATGCTGGGTTCC 9693
992 QY TGTGTTTCATATCCATTTCTGTTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATTGAGT 1051
9692 Db TGCTTACATCCAGTCTGTAGTCTATGCTCTTTTGTAGGAATTA-----GTC 9642
1052 QY CATTGATTTGAGAAATTAATCAATGACGATGTTTGTGGATTTCTGTTTA-TCTTGCATTTG 1110
9641 Db CATTGATTTGAGAGATA-TAAGGAAAGTGAATTTGTTTACTTCTGTTTATTTTGTGTTG 9582
1111 QY TGAAGT 1170
9581 Db TTAGAGTGAATTAATTTTGT 9537
1171 QY GT 1230
9536 Db ----- 9537
1231 QY CTGGAATTTATTTATTTATTTCTTGAATTTGTTGAGTGAACATCTTTAGATTTAGGTTT 1290
9536 Db -TTGAAAGATTTACTTTCTTGTGTTTCTTCTAGGTTGATTTTCCCTCTTCTGTTGTGTTT 9478
1291 QY TTCTCCTA--GCCTTCTTTAGTCTGCAATTTGAAGATAGATATTTCTTACATCTGATTTT 1348
9477 Db TCCATCTATTATCTTTTGTAGAGCTGGAATTTGTGGAATGATATTTGTGTAATTTGTTT 9418


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 16442 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-208

```

Query Match 14.2%; Score 323.899; DB 1; Length 16442;
Best Local Similarity 64.1%; Pred. NO. 8.5;
Matches 588; Conservative 0; Mismatches 296; Indels 33; G

QY	1232	TGGAAATATTATTAATTCATATTTCTTCTGAATGTGGGTAAACAATCTTTAGATTTGAAGTTTTT	1291
Db	28	TTGAAGGAATTAATTTCTATATTTTCTAGGGCGTGGTTCTATCTCTGTATGGGTTTTT	87
QY	1292	TCTCCTAGCCT-----TCTTTAGGTCTGCAATTTGAAGATAGATATTTCTTTACATC	1341
Db	88	TTTTTTTTTCTGGTTATTAATCCTTTGAAGGGCTGGATTCGTGGAGAGATAATGTGTAATTT	147
QY	1342	TGATTTTATCTTAGAATGTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTTTCTTAAGT	1401
Db	148	TGGTATTGTGATGGAATTAATTTGTTTTCTCAATCATATGGAATTTGAGAGATTTGGTTGGGT	207
QY	1402	GCAGTAGCTGGCGTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGGGCC	1461
Db	208	ATAGTACCCCTGGGTGGCGTTGGTTCTCTTAGGGTCTTTATACACATCTGTCTAGGATC	267
QY	1462	TTCTTACATTTTCAGTTTCTTATTGGAAAAGTCAGGTGTAATTTCTAAATACATCTGCCCTTA	1521
Db	268	TTCTGGGCTTTCATAGTCTCGGTGCCAAGGTCTGGTATAATTTCTGAAGCCCTGCCCTTA	327
QY	1522	TATGTTAAATGGTCTTTTTCCTTTCCCTTGCACTCTTTTAATTAATTTCTTTTGTTCATACCTT	1581
Db	328	TATGTTACTT-GACTTTTTTCCCTTACTGCTTTTAATTAATTTCTATCTTTATTTAGTGCAT	386
QY	1582	TAGTGATTTGCATTAATATGCACCTGTGGGAGTTTCTTTCCGGTCCCAATCTATTTGTGT	1641
Db	387	TGTTGTTCTGATTAATTAAGTGTGGGAGGAATTTCTTTCTGGTCTCTCTATTTGGAGT	446
QY	1642	TTTGATGTCTTGTACCTTGATAGGCATCTCTTCTCAAGGTAGSAAATTTTCTCTTT	1701
Db	447	TCGTAGGCTTCTGTATGTTCTATGTTCTATGCACTCTCTT---AAGTTTGGGAAGGTTTCTTTC	502
QY	1702	TTTGGTTTTCTTGAAATATTTTCCCTGCTTTTGACCTG-CCTTCTTCCCTTCCCTCTAT	1760
Db	503	TATTAATTTGTGGAAGATAATTGTGGCCCTTTAAGTTGAAAATCTTTCAATTTCACTAC	562
QY	1761	TCC-----TTTGGTTTTGCATAGTCTCTGGCTCTCCCTGGATGTTTTATGC	1807
Db	563	TCCTATTATCCGTANGTTTGACTTCTCATGTGTCCTGAAATTCCTCGAATTTTAAGT	622
QY	1808	CTGATTAATTTTAGACTTAACATTTTCTTTGACCAAGGTATCAATTTCTCTATCTTGTGTC	1867
Db	623	TAGGATCTTTTGTCAATTTTGCAATTTCTTTGATTTGTGCTGCTATGTTCTCTATGGAATC	682
QY	1868	TTCACTGCCCTGAGATTCCTCTTCTATCTCTGTATCTCTGAGGAGCTGTCTCTGA	1927
Db	683	TTCTGCACCTGAGATTCCTCTTCCATGTCTGTATTTCTGCTGCTGAATGCTTGCACTAT	742
QY	1928	GGTTCCTG-TTGGGTTCTTAATTTTTTCAATTTCCAGATTTTCCTTCAGTTTGGGTTTTGTT	1986
Db	743	GGTTCAGATTTCTTTTCTAGGGTTTCTATCTCTAGCGTTGGCTCAATTTTGGGTTTTCTT	802
QY	1987	TATTAATCTAATTTCCACTTTCAGGTCCGGAATGTTTTTACTCATTTTTCTCCAGTATT	2046
Db	803	TAATGTGTCTACTTCGCTTTTTAGTCTACTATGAGTTTGTGTCAATTTCCATCACCTAAT	862
QY	2047	---TACATTTTCATAGGTTTCTTTAAATGGATTTATTCATTTCTCTTCAAGGACCTTTTA	2103
Db	863	GGATGTGTTTTCTGTTTTTCTTTAAGACTTCTACCTGTTGGTTATTTTTTTCGTGTTT	922
QY	2104	TGAATTCATAAAATGTA	2120

Db 923 TTCTTAAAGGACTTGTA 939

RESULT 8

US-09-618-166-208

; Sequence 208, Application US/09618166

; Patent No. 6583112

; GENERAL INFORMATION:

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO WERNER'S SYNDROME

NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle
STATE: Washington

COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM
MEDIUM TYPE: FLO

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000

CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: McMasters, David D.
REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 240052.419C1
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:

LENGTH: 16442 base
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear

SEQUENCE DESCRIPTION; SEQ ID NO: 208:
8-166-208

Match	Score	DB 1	Length
14.2%	323.899	DB 1	16442
54.1%	Bred	NO	8.5
Local similarity			

Local Similarity 64.1%; Pred. NO. 8.5;
s 588; Conservative 0; Mismatches 296; Indels 33; Gaps 7;

1232 TGGAAATTATTATTTCATATTTTCTTGAATGTTGGGTAACTCTTAGATTGAAGTTTT 1291

28 TTGAAGGATTACTTCTTATTTTTTCTAGGGCGTGGTTCTATCCTTGTATGGGTTTTT 87

1292 TCCTCTAGCCT-----TCTTTAGGCTCGCATTGAAGATAGATATCTTTACATC 1341

88 TTTTTCCTGTTATTATCCTTTGAAGGCTGGATTTCGTGGAGAGATAATGCTGAATT 147

1342 TGATTTTATCTTAGAATGCTTTCTCTTTCTCCAACTATTGTGACAGAAAGTTTTTCTTAAGT 1401

148 TGGTATTGTCATGGAATACTTTGTTTTCTCCATCTATGGCAATTGAGAGTTTGGTTGGGT 207

1402 GCAGTAGTCTGGCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGACGGGCC 1461

208 ATAGTAGCCTGGGCTGGCGTTTGTTCTCTTAGGCTCTTTAACAATCTGCTAGGATC 267

1462 TTCTTACATTTTGAGTTTCTATTGGAAAAAGTCAGGTGTAATTCATAATCATCTGCCTTTA 1521

268 TTCTGGCTTTCATAGTCTCTGGTGC A AAGGTCTGGTATAATCTGATAGGCTGCCTTTA 327

1522 TATGTTAATTGGTCTTTTTCCTGCATCTTTTAAATATTCTTCTTGTCTATACTTT 1581

```
Db 328 TATGTACTT-GACTTTTTCCCTTACTGCTTTTAAATATCTATCTTATTTAGTGCAC 386
Qy 1582 TAGTCATTGATTATTATGACCTGTGGGAGTCTTTTTCGGTCCCAATCTATTTGGTGT 1641
Db 387 TGTGTTCTGATTATATGTTGGGAGGAATTTCTTTCTGGTCTGCTATTTGGAGT 446
Qy 1642 TTGTATGCTCTTGATACCTTGATAGGACATCTTTCTCAAGGTAGAAATTTTCTTT 1701
Db 447 TCTGTAGGCTTCTGTATGTTTCATGTGATCTCTTT-----AAGTTTGGGAAGTTTCTTC 502
Qy 1702 TTGTGTTTCTGAAAATATTTTCCCTGCTTTTGACCTG-CCTTCTTCCCTTCTCTAT 1760
Db 503 TATTAATTTGTGAAGATATTTTGTGGCCCTTTAAGTTGAAAATCTTCATTTTCATCTAC 562
Qy 1761 TCC-----TTTGGTTTTGCATAGTGTCTCTGGCTTCTCTGGATTTTATGTC 1807
Db 563 TCCATATATCCGTANGTTTGGACTTCTCATTTGTGCTGAAATTTCTCTGGATGTTTAAAT 622
Qy 1808 CTGGATTAATTTAGACTTAAATTTTCTTTTGACCAAGTATCCATTTCTCTATCTTGTGTC 1867
Db 623 TAGGATCTTTTGTGCAATTTTCTTTCATGTTTCTGATGTTGCTGATGTTCTATGGAATC 682
Qy 1868 TTCACGTGCTGAGATCTCTCTCTATCTCTTCTGTTATCTGTGACGTGAGGCTTGTCTCTGA 1927
Db 693 TTCTGCACCTGAGATCTCTCTTCCATGCTTGTATCTGTATCTGCTGATGCTTGCATCTAT 742
Qy 1928 GGTTCCTG-TTGGGTTCTTAATTTTTCATTTCCAGATTTCCCTTCAGTTTGGGTTTGT 1986
Db 743 GGTTCAGATTTCTTTCTCTAGGTTTCTATCTAGCGTTGCTTCCATTTGGGTTTCTT 802
Qy 1987 TATTAATTTCTATTTCCACTTTTCAGTCTGAAATGTTTACTCATTTTCTCCCTCAGTAT 2046
Db 803 TATGTGCTACTTCTCGCTTTTAGTCTACTATGTTTGTCTCAATTTCCATCAGTATTT 862
Qy 2047 ---TACATTTTCATAGTTTCTTTAAGGATTTATTCATTTCTCTTCAAGGACCTTTTA 2103
Db 863 GGATGTTTCTCTGTTTCTTTTAAAGGACTCTACCTGTTTGGTATTTTTCGTGTTT 922
Qy 2104 TGAATTCATAAATGTA 2120
Db 923 TTTCTTAAGGACTTGTGA 939

RESULT 9
US-09-497-855A-37
; Sequence 37, Application US/09497855A
; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UM01523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatenIn version 3.0
; SEQ ID NO 37
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-37

Query Match 11.1%; Score 253.097; DB 1; Length 193303;
Best Local Similarity 51.9%; Pred. No. 1.7;
Matches 861; Conservative 0; Mismatches 684; Indels 115; Gaps 14;

Qy 292 TCTATTCTTGATTTCTATCTTGGCTCAATTTTAACTCAGTAGTGAAGTTTCTGCTTCC 351
Db 110168 TTTCTGCTTTCATTTCAATATGATACCCAGTAGTCAATTCAGGAGCAGGTTGTTCACTTCC 110227
```

```
Qy 352 ATAAGTTTGTAAAGTTTCTGTTGTTTCTGTTGTTTCTGTTGTTTCTAGATTTAAGCTGT 411
Db 110228 ATGTAGTTCTGAGTGGTTTGTGTGAGT-----TTCTTAATCCTGAGTTCTAGTTGATGAC 110283
Qy 412 GGTGGTTCAGATAGGACATAGAGTATTTATTTCAATTTGCTTTTATCTGTCGAGACTTGCTT 471
Db 110284 TGTGGTCTCAGAGACAGTTTGTATAATTTCTGGTCTTTCACATTTGCTGAGAGTGTCTT 110343
Qy 472 TGTTTTGAATATGTAATCAATTTTGA--GAGTTTTCATAGGGTCTGACAAAGAGTAC 529
Db 110344 TACTTCCAACTATGTGGTCAATTTTGAATAGGTGTGGTGTGGTGTGAAAGAAATGAT 110403
Qy 530 AGTCT--TTGTGTTTGGTGAAATAGTCTGTAATATCTCTAGTCCACTTGTGTTATGAC 588
Db 110404 ATTCTGTTGATTTGGGTGGAGAGTTCTGTAGATGCTATTTAGGTCTGCTTGGTCAGAG 110463
Qy 589 ATCAGTTAGTCTCCAGCAATTTCTCTGTTTCTGTTTCTGTTTGTGAGATGACCTAATCTTGGAG 648
Db 110464 CTGAGTTCAATTCCTGGGTGTCCTTTGTTAGCTTTCTGCTTCTGATCTGTCTAATGTTG 110523
Qy 649 AGATCGGGTATTGAAGTAGCCCACTACTGCTGTGTGAG-GTCAATATGTGATTTTAGCT 707
Db 110524 ACAGTGGGTGTAAAGTCTCCCAATTTTATTTGTGTGGAGTCTAAGTCTCTTTGTAGT 110583
Qy 708 GTAGCTGTCTTGTGTTTATGAACCTTGGGTGACATTTGTGTTTGGTGCATAGACATTAAGAA 767
Db 110584 CACTCAGGACTTGTCTTATGAATCTGGGTGCTCTGTATTGAGTGCATATATATTAGGA 110643
Qy 768 TTGCAATGCTCTCTGGTGGATTTTCTTTTGANGCCTATGTAATTTCTTCCCAATCTCA 827
Db 110644 TAGTTAGCTCTTCTTGTGTAATGATCCCTTACCATTATGTAATGGCCCTTTTTCCTCT 110703
Qy 828 TCTGCTTAGTTTGGGTTTAAAGTCT-ATTAGTCAGATATTAATAAGTGTATCGGCTTG 886
Db 110704 TTTGATCTTTGTGTTTAAAGTCTGTTTATCAGAGACTAGATTTGCAACCTTGCCTT 110763
Qy 887 CTTCTTGGGCAATTTGCTTAGAATATCTT-TTCCATCCTTTTACTCTAAGTGAATGCT 945
Db 110764 TTTTGTGTTTCCATTTTCTTGGTAGATCTTCTCCATCCCTTATTTTGGAGCTATGTGT 110823
Qy 946 ATCCAT---GGTAGTTGCTTTTGGATGAGCAGTAGGATGATCTTGTGTTTCATAT 1002
Db 110824 GTCTGCGAGTGGAGTGGTTTCCCTGAATACAGCACACTGANGGCTTGTACTCTTTAT 110883
Qy 1003 CCATTTCTGTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATTGAGTCATTGATGTTG 1062
Db 110884 CCATTTGCCAGTCTGTGCTTTTAAATGGAGCAATTAGCCCAAT----- 110928
Qy 1063 AGAATTATCAATGACAGAGTTTGTGGATCTTGTATCTTGCACCTTGAAGTGTGTGT 1122
Db 110929 -----TACATTTAAGGTTGATATTGTTATGTTGTAAT 110960
Qy 1123 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1182
Db 110961 TTGATCCTGTGTTATG-----AGTTAGTGTGTCATTTT 110995
Qy 1183 GTCGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1242
Db 110996 GCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 111055
Qy 1243 ATTATTCATATTTTCTTGAATGGGTGAACATCTTTAGATTGAAGTTTTCCTCCTAGCT 1302
Db 111056 TACAGTGGCTGGTATTTGGTGTGTTTCTTTCCATGTTTAG-----TGTCTTCTCAGAGCT 111110
Qy 1303 TCTTTAGGCTGCATTTGAAGATAGATATTTCTTTACATCTGATTTTATCTTGAATGTCT 1362
Db 111111 CTTTTAGGCGAGTCTGTGGTGAACAAATCTCTCAGCAATTTGCTGTCTGTAAGTAT 111170
Qy 1363 TTTCTTCTCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCTCAGATC 1422
Db 111171 TTAATTTCTCCTCACTTATGAAGCTTAGTTTGGCTGGATATGAAATTTCTGGTTGAAAT 111230
```



```

; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: 6510
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "D = A or G or T"
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: 3406..6470
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "K = G or T"
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: 3564..7896
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "M = A or C"
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: 3497..3607
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "N = A or C or G or T"
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: 3479..6422
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "R = A or G"
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: 3405..6871
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "S = C or G"
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: 3457..9998
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "W = A or T"
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: 3595..9999
; OTHER INFORMATION: /mod_base= OTHER
; OTHER INFORMATION: /note= "Y = C or T"
; US-08-772-440-33

Query Match 11.1%; Score 251.899; DB 1; Length 10409;
Best Local Similarity 68.0%; Pred. No. 30;
Matches 463; Conservative 5; Mismatches 176; Indels 37; Gaps 9;

QY 1385 AGAAGTTTTTCTAAGTGCAGTAGTCTGGCCTGCATCTGPACTCTTTGGAGTCTGTAG 1444
DB 4097 AAAATAATTAATCTGGCTAGCAGTCTGGGCTCATATCTGTGTTCTTAGAGTTTGTAG 4156
QY 1445 CACATCTGTGAGGCGCTTCTACATTTGAGTTTCTATTGGAAAGTCAAGTCTAATTC 1504
DB 4157 AACATCTGTTGAGGCGCTTCTGAGTTTCTAGTCTTCAAT-GAGAGTCAAGGCAATATTC 4215
QY 1505 TAATACATCTGCTTTATATGTTAATTTGCTTTTCCCTTGCATCTTTAATATTCCTT 1564
DB 4216 TAGTAGGTTTATATATATGTTACTCAGTC-TTTTCTCTTGCAGCTTTTCATATCTT 4274
QY 1565 TCTTGTCTTACTTTTATGATTTGATTTATATGACTGTGGGAGTCTTTTCGG 1624
DB 4275 TCTTGTCTGATGTTTATGATTTGATTTATGTTGTTGGGGAATATCTTTCTGA 4334
QY 1625 TCCATCATTTTGGTGTGTTTGTATGCTTCTGTACCTTGATGAGCACTCTTTCTCAAG 1684
DB 4335 TCCAGTCCATTTGGTGTCTGATGTTTGTACCTTGATTAATCACTC---CTTAAGG 4391
QY 1685 TTAGGAATTTTCTTTTGTGTTTCTTGAATAATTTTCCCTGCTTTTGACCT--GCC 1742
DB 4392 TTGAAGACATTTCTTTTATGATTTGTTGAAATAATTTCTGTGCAATTTTAACTTGGCY 4451
QY 1743 TCTTCCCTTCCCTTATCTTCTTTGG-----TTTTTGCATAGTGTCTC 1785

```

4452 TTCTTTTCTTTGCTATATCCVATAGACTAGTTTGTGTTGTTTGTGTTTAACTG 4511

1786 TGGCTTCTCGATGTTTATGCC---TGGATTATTTAGACTTAACATTTCTTTGAC- 1840

4512 AGAATTCCTGGATGTTTGTGCCCTKGAGTTTTTTTTTTTWTATTAATATTTCTKGACT 4571

1841 CAAGGTATCCATTTCTTCTATCTTCTTCACTGCTGAGATTTCTCTCTCTAATCTTCTG 1900

4572 GGAGATAYWCCTTTCTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTA 4631

1901 TATT---CTGTCAGTGAGGCTTGTCTCTGAGTTCTCTGAGTTCTCTGTTGGTTCTTAATTTTCAAT 1957

4632 TACTGAGTTGTTGGTGAGGCACTCTTCTAGACCTTGTGTTGTTCTCCCAATATATACAT 4691

1958 TCCAGATTTCTTCACTGTTGGGTTTGTGTTTATTAATTTCTATTTCCATTTTCAGGTCCTGA 2017

4692 TCCAG-----TTTCATGCTAGTTTTTACTTATCGATTATATTTTACTTTTCATGACATAA 4746

2018 AATGTTTACTCATTTTCTC 2038

4747 CTCTTTTCAATATTTTATTC 4767

RESULT 13

US-09-426-290-1

; Sequence 1, Application US/09426290

; Patent No. 6410712

; GENERAL INFORMATION:

; APPLICANT: Berglind Ran Olafsdottir

; APPLICANT: Jeffrey Gulcher

; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE

; FILE REFERENCE: 2345, 2001-000

; CURRENT APPLICATION NUMBER: US/09/426,290

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 168575

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (21181)...(21403)

; NAME/KEY: CDS

; LOCATION: (95252)...(95430)

; NAME/KEY: CDS

; LOCATION: (101753)...(101996)

; NAME/KEY: CDS

; LOCATION: (110324)...(110439)

; NAME/KEY: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; US-09-426-290-1

Query Match 10.9%; Score 248.997; DB 1; Length 168575;

Best Local Similarity 51.8%; Pred. No. 2;

Matches 857; Conservative 0; Mismatches 695; Indels 102; Gaps 14;

QY 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTTTAACACTGAGTAGTGAGTTGTTGTTTCC 351

DB 156020 TTTTGCCTTCAATTTTGTATGATACCCAGTAGTCAATCAGGACAGGTTGTTTCC 156079

QY 352 ATAAGTTCTGATGTTTCTGTTGTTTCTGTTGTTTCTGTTGTTTCTAGATTAAAGCTGT 411

DB 156080 ATGAGTTGAGGAGTTTGTGAGTACT---TTCCTAATCCTGAGTTCTAGTTGATGAC 156136

QY 412 GGTGTCAGATAGGACATAGATATTTTCAATTTGCTTTTATCTGTCGAGACTTGTCT 471

DB 156137 CGTGGTCTGAGACAGATTTGTTATATATCTGATCTTATACATTTGCTGAGGAGCTT 156196

```
QY 472 TGTGTTGAAATATGTAATCAATTTTGGG--GAGTTTCATAGGGTGCTGACAGAGGTAC 529
Db 156197 TACTTCCAACTATGTGGTCAATTTTGGAAATAGGTGTGGTGCTGCTGAGAGAAATGAT 156256
QY 530 AGTCT-TTGCTGTTGGTGAATAGTCTGTGAATATCTCTAGGTCCACTTGGTTTATGAC 588
Db 156257 ATTCTGTGTAATTCGGGTGGAGAGTCTGTAGATGTCATTAAGTCTGCTGGTGCAGAG 156316
QY 589 ATCAGTTAGTCCAGCAATTCCTGTGTTTCGTTTGTGTTGATGATGACTCACTGTTGGAG 648
Db 156317 CTGAGTTCAATTCCTGGATATCCCTGTGTTAACTTCTGTCTGTTGATCTGCTTATATTG 156376
QY 649 AGAATGGGGTATGAAGTAGCCCACTATCTGTGTGT--GAGTCAATATGTGATTTAGCT 707
Db 156377 ACAGTGGGGTGTAAAGTCTCCCAATTAATTAATGTTGTGTAGTCTAAGTCTCTTTGTAGT 156436
QY 708 GTAGCTGTCTGTTGTTTATGAACCTGGGTGACATGTTGTTGGTGATAGACATTAAGAA 767
Db 156437 CACTCAGGACTGTCTTTATGAATCTGGGTGCTCCTATATTTGGGTGCATATATATTTAGGA 156496
QY 768 TTGCAATGCTCTCTTGGTGGATTT--TCCTTTGATGCTATGTAGTATTTCTCCCAATCTC 826
Db 156497 TAGTTAGCTCTCTGTTTCAATTAATCCCTTACATTAATGTAATGGCCTTCTTTGTC 156556
QY 827 ATCTGCTTGTGTTTGGGTTTAAAGTCT--ATTAGTCAGATATTAATAAGTACTGTATCGGCTT 885
Db 156557 TTTTGAATCTTGTGTTTAAAGTCTGTTTATCAGAGACTAGGATTCGAACCCCTGCCT 156616
QY 886 GCTTCTTAGGGCCATTTGCTTAGATATCTTT--TCCATCTCTTTTACTCTAAGTGATGTC 944
Db 156617 TTTTGTGTTTCCATTTGCTTTGGTAGATCTTCCCTCCATCTCTTTTACTTTGAGCCTATGTG 156676
QY 945 TATCCATG--GTAGGTTGCTTTTGGATGACAGCAGTAGGATGATCTTTGTTTTCATA 1001
Db 156677 TGTCTGCACTGAGATGGTCTCCTGAATACAGCACACTGATGGTCTTTGACTCTTTA 156736
QY 1002 TCCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATGTT 1061
Db 156737 TCCAATTTGCCAGTCTGTGCTTTTAAATGAGCAATTAGTCCCTT-----156782
QY 1062 GAGAATTATCAATGAGCAGTGTGTTGTGATCTCTGTTTATCTGTGCACTTGTGAAGTGTCG 1121
Db 156783 -----TACATTTAAAGTTAATA 156799
QY 1122 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1181
Db 156800 TTGTTATGTGTGAATTTGATCCTGCTCATTTGA-----ATGTTAGCTGCTTATTT 156848
QY 1182 TGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1241
Db 156849 TGTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156908
QY 1242 TATTATTCATATTTTCTTGAAATGTGGGTAAACATCTTTTAGATTTGAAGTTTTCCTAGACC 1301
Db 156909 TTGCAGTGGCTGTACTGTTGTTCCCTTCCATGTTAG-----TGCTTCTTCAGAGC 156963
QY 1302 TCTTTTAGTCTGCAATTTGAAGATAGATATTTCTTTTACATCTGATTTTATCTTGAAGTCT 1361
Db 156964 TCTTTTAGGGCAGGCCTAGTGTGTGACAAAATTTCTCAGCAATTTGCTGTCTGTAAGGAT 157023
QY 1362 TTTCTTTCTCCAACTATTTGTACAGAAAGTTTTCCTAAGTGCAGTGTCTGGCTGACAT 1421
Db 157024 TTTATTTCTCTTCACTTATGAGCTTAGTTTGGCTGATATGAAATTCGGTTGAAA 157083
QY 1422 CTGTAGTCTCTTGTAGTCTGTAGCACATCTGTGTGAGGGCCCTCTTACATTTTGTAGTTTCT 1481
Db 157084 TTTCTTTCTTTTAAAGATGTTGAATTTGGGCCCCACCTCTCTTCTGACTGTGTAGAGTTTCT 157143
QY 1482 ATTGGAAGTCAAGTGTAAATCTTAATACATCTGCCCTTATATGTTAATTTGGCTTTT 1541
Db 157144 GCCGAGAGA-TCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 157201
QY 1542 CCTTGCATCTTTTAAATCTCTTTTGTGTTCTATACCTTTTAGTGAATTTGATTATTATGC 1601
```

RESULT 14

US-09-818-512-3

; Sequence 3, Application US/09818512

; Patent No. 6537780

; GENERAL INFORMATION:

; APPLICANT: BEASLEY, Ellen et al.

; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES

; FILE REFERENCE: THEREOF

; FILE REFERENCE: CL001192

; CURRENT FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 116592

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)...(116592)

; OTHER INFORMATION: n = A,T,C or G

US-09-818-512-3

Query Match

Best Local Similarity 10.7%; Score 242.997; DB 1; Length 116592;

Matches 847; Conservative 0; Mismatches 705; Indels 102; Gaps 13;

```
QY 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTAGTAGTAGTGTGTTGTTTCC 351
Db 51756 TTTCTGCCCTTCAATTCGTTTATGTACCCAGTACTCATTGAGGAGCAGGTTGTTTCC 51815
QY 352 ATAAGTTTCTAAAGTTTCTGTTGTTTCTGTTGTTGTTTATCTAGATTTAAGCTGT 411
Db 51816 ATGTAGTTGAGCG---GCTTTGAGTGAGATTCCTTAATCCTGAGTTCTAGTTGATGCAC 51872
QY 412 GGTGTTGAGATGAGACATAGAGTATTTTCAATTTGCTTTTATCTGTGAGAGTTGCTT 471
Db 51873 TGTGTTCTCAGAGACAGATTTGTTATAAATTTCTGTTCTTTTACATTTTCTGAGGAGCTT 51932
QY 472 TGTGTTGAAATATGATCAATTTTGGAA--GAGTTTCATAGGTCCTGACAGAGGTAC 529
Db 51933 TACTTCCAACTATGTGTTCAATTTTGGAAATAGGTGTGTTGTTGTTGTTGTTGTTGTTGTT 51992
QY 530 AGTCT-TTGTGTTTGGTGAATAGTCTGTAAATATCTTAGGTCACCTTGTGTTTATGAC 588
```

Db 51993 ATTCGTGTTGATTTGGGGTGGAGAGTCTCTGTAGATGTCATATAGTCTGCTTGGTCAGAG 52052
QY 589 ATCAGTTAGCTCCAGCATTTCTCTGTTTCGTTTCTGTTGAGATGACCTAACTGTTGGAG 648
Db 52053 CTGAGTTCAATTCCTGGGTATCCCTTGTGTGACTTTCTGTCTGTTGATCTCTGTCTAAATG 52112
QY 649 AGAATGGGGTATTGAAGTAGCCCACTATCTGTGTGTGAG-GTCAATATGTGATTTTAGCT 707
Db 52113 ACAGTGGGGTGTAAAGCTTCCATTAATATATGTTGGAGTCTAAGTCTCTTGTAGGT 52172
QY 708 GTAGCTGTCTGTTTATGAACTTGGGTGACATGTGTTTGGTGATAGACATTAAGAA 767
Db 52173 CACTGAGGACTTGCTTTATGAATCTGGGTGCTCTGTATTTGGGTGCAATAATATTAGGA 52232
QY 768 TTGCAATGCTCTCTTGGTGGATTT-TCCCTTGTGATGCTATGTAGTATCTTCCCAATCTC 826
Db 52233 TAGTTAGCTCTCTTGTGAAATGATCCCTTACATTAATGTAATGAGCTTCTTGTCTC 52292
QY 827 ATCTGCTTGTGTTTGGGTTTAAAGTCT-ATTAGTCAGATATTAATAAGCATGTATCGGCTT 885
Db 52293 TTTTGATCTTTGTTGTTTAAAGCTGTTTATCAGAGACTAGGATTTGCAACCCCTGCCT 52352
QY 886 GCTTCTTAGGGCCATTTGCTTAGATATCTT-TTCCATCTCTTTTACTCTAAGGTGATGC 944
Db 52353 TTTTGTGTTTCCATTTGCTTGGTAGATCTTCCCTCATCTCTTTTATTTTGGAGCTATGTG 52412
QY 945 TATGCATG---GTAGTTGTCTTTTGGATGACAGTAGTAGGATCTTGTGTTTTCATA 1001
Db 52413 TGTCTCTGACGTGAGATGGGTTTCCGAAATACACACACTGATGGGCTTGTACTCTTTA 52472
QY 1002 TCCATCTGTATCCAGCATCTTTTCTAGAGAAATTAAGATCATTCAGTATGATGTT 1061
Db 52473 TCCACCTTGCCAGCTGTGTCTTTTAAATGCAAAATTTAG-----TCCATTTATATT 52524
QY 1062 GAGAAATATCAATGACGAGCTGTTTGGATCTGTTTATCTTGCATCTTGCAAGTGTGTG 1121
Db 52525 TAAAGTTAAATGTTATGTTGTAATTTGATCTCTGTCATTAATGATGTAGTGTGTGATTT 52584
QY 1122 TGT 1181
Db 52585 TGCTCATTAGTTGATGCAATTTCTTCTAGTCTGAGGCTTTCATATTTTGGCATGAT 52644
QY 1182 TGTCTGT 1241
Db 52645 TTGAGCGGCTGGTACCGGTTGTTCCCTTTC----- 52675
QY 1242 TATTATTCATATTTTCTTGAATGTGGGTAAACATCTTTAGATTTGAAGTTTTCCTTAGCC 1301
Db 52676 -----ATGTTTAGCGCTTCTCTCAGGAGC 52699
QY 1302 TCTTTTAGGCTGCAATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATGTC 1361
Db 52700 TCTTTTAGGGAGGCGCTGGTGGTGACAAAATCTCTCAGCAATTTGCTGTCTATAAGTAT 52759
QY 1362 TTTCTTTCTCAACATTTGTCACAGAAAGTTTCTTAAGTGCAGTAGCTGCGCTCGACAT 1421
Db 52760 TTTATTTCTCTTCTTACATGAGCTTAGTTTGGCTGGATAGAAATCTGGGTGGAAGA 52819
QY 1422 CTGTAGTCTTTGGAGTCTGTAGACATCTGTGTGAGGGCCCTTCTTACATTTTGTGTTTCT 1481
Db 52820 TTTCTTTCTTAAAGAAATGTTGAATATTGGCCCCCACTCTCTCTCTGGGCTTGTAGGGTTCT 52879
QY 1482 ATTGAAAGTCAGGTGTAATTTCTAAATACATCTGCTTTTATCTTAAATGTTGTTCTTTT 1541
Db 52880 GCCGAGAG-TCCGCTGTAGTCTGATGGGCTTTTCTTTG-AGGGTAACCCGACCTTTCT 52937
QY 1542 CCGTTGCACTTTTAATATCTTTCTTTTGTGTTTACTTTTGTAGTATTTGATTTATGTC 1601
Db 52938 CTCCTGCTGCCCTTAAACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 52997
QY 1602 ACTGTGGGAGTTTCTTTTCCGTTCCATCTTATTTGGTGTGTTTGTATGCTCTTGTACTT 1661

Db 52998 GTCTGGAGTTCCTCTCTCGAGGAGTATCTTTGTGGCGTTCTCTGTATTTCCCTGAATCT 53057
QY 1662 TGATAGGCATCTCTTCTCAAGTTTAGAAATTTTCTTTTGTGTTTCTTTTGTGAAATAT 1721
Db 53058 GAACGTTGGCCTGCTTGTAGATTGGGAAGTTCTCTCGATATATCTCTGCAGAGTGT 53117
QY 1722 TTTCC-----CTGCTTTTACCTGCTTCTTCCCTTCTTCCCTTCTTATTTCCCTTTG 1767
Db 53118 TTTCCAACTTGGTTCCATTTCTCCACATCACTTTTCAGGTACCAATCAGAGCTAGATTG 53177
QY 1768 GTTTTGTGATAGTCTCTGGCTTCTCGATGTTTTATGCTTGGATTATTTAGACTTAA 1827
Db 53178 GTCTTTTTCACATGAGTCCCAATTTCTTGAGGCTTTGCTCATTTCTTTTATCTTTT 53237
QY 1828 CATTTTCTTTTGACCAAGGTATCCATTTCTTCTATCTTCTTCACTGCTGAGATCTCT 1887
Db 53238 CTTCAAACTTCCCTTCTGCTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 53297
QY 1888 CTTCTATCTTGTATTTCTGTCTGAGTGGCTTGT 1921
Db 53298 CTTCCAGTTGATCGCATCGGCTCTCTGAGGCTTCT 53331
RESULT 15
US-09-751-389-3
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01067
; CURRENT APPLICATION NUMBER: US/09/751,389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3
Query Match 10.5%; Score 239.897; DB 1; Length 786431;
Best Local Similarity 51.4%; Pred. No. 0.51;
Matches 833; Conservative 0; Mismatches 686; Indels 103; Gaps 13;
QY 324 TAATCAGTAGGAGTGTGTTGGTTTCCATAAGTTTGTAAAGTTTCTGTTTCTGTTG 383
Db 580096 TCATTACAGGAGGAGTGTGTTTCAAGTTTCCATGATGTTTGGGTTTTCAGTGTAGT-----TTC 580151
QY 384 TTGTTGTTCTTATCTAGATTTTAAAGCTGTGGTGCAGTAGGATGATGATTTTCA 443
Db 580152 TTAATCTCTGAGCTCTAGTTTGTATCTGTTGTTCTGAGACAGTGTGTTGTAATTTCT 580211
QY 444 ATGTCTTTTATCTGTCAGAGCTTCTTGTGTTTGAATATGATTTCAATTTGGA--GA 501
Db 580212 ATTCTTTTACATTTCTGCTGAGGAGTCTTTACTTCAACTGTGTGGTCAATTTTGAATAG 580271
QY 502 GTTTCATAGGGTGTGCAAGAAAGGTACAGTCT-TTGTGTTTGTGTAATAGTCTGTA 560
Db 580272 GTGCTGTGTGTGCTGAAAGAAATGATATCTCTGTTGATTTGGGCTGGAGAGTCTGTAG 580331
QY 561 ATATCTTAGGTCCTAGTTTATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 620
Db 580332 ATGTCTTATTAGTCCACTTGTGTGAGAGTGTGATTTCAATTTCTGGGTATCTTGTAACT 580391
QY 621 TTTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 680

Db 15727 ACAGTGGGGTAAAGTCCCAATTAATTAATGTGTGGAGCTAAAGTCTCTTTGTAGGT 15786
QY 708 GTAGCTGTCTGTTTATGAACCTTGGGTGACATGTGTTGGTGCATAGACATTAAGAA 767
Db 15787 CACTGAGGACTTGTCTTATGAATCTGGGTGCTCTGTATTTGGGTGCATAAATATTTAGGA 15846
QY 768 TTGCAATGCTCTTGGTGGATT-TCCCTTGTATGCTCTATGTAGTATCTTCCCAATCTC 826
Db 15847 TAGTTAGCTCTCTTGTGAAATGATCCCTTACCAATATGTAATGGCTCTCTTTGTCTC 15906
QY 827 ATCTGCTAGTTTGGGTTAAGTCT-ATTAGTCAGATTAATAAGATGATCTGCTT 885
Db 15907 TTTTGTCTTGTGGTTTAAAGTCTGTATATCAGACACTAGGATGCAACCCCTGCT 15966
QY 886 GCTTCTTAGGGCCATTTGCTTAGAATATCTT-TTCCATCTCTTTTACTCTAAGGTGATGTC 944
Db 15967 TTTTGTCTTCCATTTGGCTGTAGACTCTCTCCATCTCTTTTATTTTGGCCATATGTG 16026
QY 945 TATCCATG---GTAGTTGTCTTTTGGATCCAGCTAGATGATCTTGTTCATA 1001
Db 16027 TGTCTCTGACGTGAGTGGGTCTCTGAATACAGACACTGATGGGTCTTGACTCTTTA 16086
QY 1002 TCCATCTCTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTCAGTCAATGTATGTT 1061
Db 16087 TCCAACTTGCCAGTCTGTGTCTTTTAAATTCAGAAATTTAG-----TCCATTTATAT 16138
QY 1062 GAGAAATTAATGAGCTAGTGTGTGGAATCTTGTATCTTGCACTTGTGAAGTGTG 1121
Db 16139 TAAAGTAAATATGTATGTGGAATTTGATCCCTGTCAATTAAGATGTAGCTGGCGATTT 16198
QY 1122 TGT 1181
Db 16199 TCTCATTTAGTGTGAGTTCTTCTAGTCTGATGGTCTTTTACATTTTGGCATGAT 16258
QY 1182 TGT 1241
Db 16259 TTGCGAGCGCTGTACCGGTGTCTCTTCC-----16289
QY 1242 TATTATTCATATTTCTTGAATGTGGTAAACATCTTTAGATTTGAAGTTTCTCTCCAGCC 1301
Db 16290 -----ANGTTTACCGCTTCTCTCAGGAGC 16313
QY 1302 TTTCTTAGTCTGCAATTTGAAGATAGATATCTTTTCAATCTGATTTTATCTTAGAATGTC 1361
Db 16314 TCTTTTAGGGCAGCGCTGGTGTGACAAATCTCTCAGCAATTTGCTTGTCTATAAAGTAT 16373
QY 1362 TTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTAAGTGCAGTCTCTGGCTGACAT 1421
Db 16374 TTTATTTCTCTCACTTATGAGCTTATTTGGCTGGAATGAATCTCTGGGTGAAA 16433
QY 1422 CTGTAGTCTCTTGGAGTCTGTAGCACTCTGTGAGGCGCTCTTCAATTTGAGTTTCT 1481
Db 16434 TTTCTTTCTTTAAGAAATGTGAATATGGCCGCCACTCTCTCTGGCTTTGAGGTTCT 16493
QY 1482 ATTGAAAAGTCAAGGTGAATTTCTAATACTCTGCTTTATATGTAATTTGGTCTTTT 1541
Db 16494 GCCGAGAGA-TCCGCTGTGTAGTCTGATGGCTTTCCITG-AGGTAACCGCACTTCT 16551
QY 1542 CCGTTGCAATTTTAAATCTTCTTTGTTCTATATCTTTTGTAGTATGATTTATGTC 1601
Db 16552 CTCGGCTGCGCTTAAACATTTTCTCTCACTTTTGGTGAATCTGCAATTTATGT 16611
QY 1602 ACTGTGGGAGTTCTTTTCCGGTCCAACTCTATTTTGGTGTGTGTGTGTGTGTGTGTGT 1661
Db 16612 GTCCTTGGAGTGTCTCTCTCGAGGAGTATCTTGTGGCGTCTCTGTATTTCTGATCT 16671
QY 1662 TGATAGGATCTCTTTCTCAAGGTTAGGAAATTTTCTTTTGGTGTGTGTGTGTGTGTGTGTGT 1721
Db 16672 GAACGTTGGCTGCGCTTCTAGATTTGGGAAAGTTCTCTCGGATAAATATCTCGAGAGTGT 16731
QY 1722 TTTTCC-----CTGCTTTTGAACCTGCGCTTCTCTCCCTTCTCTATTTCTTTG 1767
Db 16732 TTTCCAACTTGTTCCATCTCCACATCATCTTTCAGGTACACCAATCAGACGTAGATTG 16791

QY 1768 GTTTTTCATAGTGTCTCTGGCTTCTCTGGATGTTTATGCTGATTTATTTAGACTTAA 1827
Db 16792 GTCTTTTCATAGTCCCATATTTCTTGGAGCTTCTGCTCATTTCTTTTATCTTTT 16851
QY 1828 CATTTCTTTTACCAAGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATCTCT 1887
Db 16852 CTCTAAACTTCCCTTCTGCTTCTATTTCAITTCATTTCAITTCATTTGCTGATACCTTT 16911
QY 1888 CTCTCATCTCTTGTATTTCTGTCAGTGGCTTGT 1921
Db 16912 CTTCAGTTGATCCATCGGCTCTCTGAGCTTCT 16945

RESULT 17

US-09-340-620A-63
; Sequence 63, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 32042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-620A-63

Query Match 10.58; Score 239.797; DB 1; Length 32042;
Best Local Similarity 51.18; Pred. No. 11;
Matches 845; Conservative 0; Mismatches 707; Indels 102; Gaps 13;
QY 292 TCTATTTCTTGATTTCTATCTTGCTCATTTTAACTCAGTGTGTTGGTTTCC 351
Db 15370 TTTCTGCTTCATTTCTGTTATGTACCCAGTAGTCAITTCAGGAGCAGGTGTTTCAGTTTCC 15429
QY 352 ATAAAGTTTGAAGTTTCTGTTCTGTTTCTGTTTGTGTTGTTTATCTAGATTTAACTGT 411
Db 15430 ATGTAGTTGAGCG---GCTTGTAGTGTAGTCTTAATCTCTGAGTTCTAGTTGTGAC 15486
QY 412 GGTGCTCAGATAGACATAGAGTATTAATTTCAATGTCTTTTATCTGTCTGAGATTTGCTT 471
Db 15487 TGTGCTCTGAGAGATAGTTTGTATTAATTTCTGTTCTTTTACATTTCTCTGAGAGACTT 15546
QY 472 TGTTTTGAATATGATTAATCAATTTTGA---GAGTTTCAAGGCTGCTGACAAAGGTAC 529
Db 15547 TACTTCCAACTATGTGTGTCATTTTGAATAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTT 15606
QY 530 AGTCT-TTGTGTTTGTGAAATAGTCTGTAATAATCTCTAGGTCCACTTGGTTTATGAC 588
Db 15607 ATTCTGTGTTGTTGGGTGGAGAGTTCTGTAGATGTCATTAGGTCTGCTGTGTCAGAG 15666
QY 589 ATCAGTTAGCTCCAGCATTTCTCTGTTTCTGTTTGTGTTTGTGTTGTTGTTGTTGTTGTTGTT 648
Db 15667 CTGAGTTCAATCTCTGGGTATCTTGTGACTTTCTGTCTCATTTGATCTGTCTAATGTTG 15726
QY 649 AGAATGGGGTATTGAAGTACCCCACTATCTGTGTGAG-GTCAATATGTGATTTTGTAGCT 707
Db 15727 ACAGTGGGGTAAAGTCTCCCAATTTAATGTGGGAGTCTAAGTCTCTTTTGTAGGT 15786
QY 708 GTAGCTGTGCTGTTTGTATGAACCTTGGGTGACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 167

Db 145434 ATATAGTTGAGCGGTTTGTAGTGAGT---TTCTTAATCCTGAGTCTAGTTGATGCA 145490
QY 412 GGTGTCAGATGAGACATAGAGTATTAATTTCAATTTGCTTTTATCTGTCGAGACTTCTCTT 471
Db 145491 TGTGTCGAGAGACAGTTGCTTAATTTCTGATCTTTTACATTTGCTGAGAGAGCTT 145550
QY 472 TGTGTTGAAATAGTATTAATTTTGA--GAGTTTCATAGGTCGTCGACAGAGGTAC 529
Db 145551 TACTTCCAACTATGTGGTCAATTTTGAATAGGTGTGGTGTGGTGTGCTGAAATAATGTAT 145610
QY 530 AGTCT-TTGTGTTTGTGTAATAGTCTGTAATATCTCTAGGTCACATTTGGTTTATGAC 588
Db 145611 ATTCTGTTGATTTGGGTCGAGAGTCTGTAATATCTCTAGGTCGCTGATGTCAGAG 145670
QY 589 ATCAGTGTAGTCCAGCAATTTCTGTTTGTGTTTGTGTTGATGAGTACCTAATGTTGGAG 648
Db 145671 CTGAGTTCAATTCCTGGGTATCTCTGCACTTTCTGCTGCTGATCTCTAATGTTG 145730
QY 649 AGAATGGGTATTCAGATGAGTACCTAATCTGTTGTGTGAG-GTCAATATGATTTTGTAGCT 707
Db 145731 ACAGTGGATGTTAAAGTCTCCCAATTAATTTTGTGGAGTCTAAGTCTCTTTGTAGGT 145790
QY 708 GTAGCTGTGCTGTTTATGAACCTGGGTGACATTTGTTGTCATAGACATTAAGAA 767
Db 145791 CACTCAGACTTGTCTTATGAATCTGGGTGCTCTGTTATGATACATATATTTAGGA 145850
QY 768 TTGCAATGCTCTCTGTG-GAATTTCTGTGATGCTATGATGATTTTCCCAATCTC 826
Db 145851 TAGTGTAGTCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 145906
QY 827 ATCTGCTTGTGTTTGGGTTTAAAGTCTATAGTACATTAATAATGATGATGCTGCTTG 886
Db 145907 TTTTGTGCTTGTGTTTAAAGTCTGTT--TTATCAGACTATGATTCGACCCCTGTC 145964
QY 887 CTTCCTTAGGCGCAATTCCTTAGAATATCTTT-----TCCATCTCTTTTACTC 932
Db 145965 CTCTTTTGTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 146024
QY 933 TAGGTGATCTCTATCCATG---GTAGTTGCTTTTGTGATGTCAGAGTAGATGAT 989
Db 146025 TGAGCTATGCTGCTCTGACGTCGTGATGATGATGATGATGATGATGATGATGATGATG 146084
QY 990 CTGTTTTTCATATCCATCTGTTACCCAGTATCTTTTCTAGAGAATTAAGATCAATGA 1049
Db 146085 CTTGACTCTTTATCCAAATTTGCCAATCTGCTCTTT-----AATTAGACATTCAG 146136
QY 1050 GTCAATGATGTCAGAAATATCAATGACGAGTGTGTTGTTGATTCCTTTTATCTTGCACTT 1109
Db 146137 CCATTTTACCTTTAAGGTTAATATGTTATGTCGAATTTGATCCTGTCATTAATGATTT 146196
QY 1110 GTGAAGTG 1169
Db 146197 AGCTGTTATTTGCTTTGTTACTTGATGAGTACTTCTTAGCATGATGCTTTTACAA 146256
QY 1170 TG 1229
Db 146257 TTTGGCATGTTTGTGAGTGGCTGACCAATGTTTCTCTCC----- 146299
QY 1230 CCTGGAATTTATTTATTTATTTCTTGAATGTTGGGTAAACATCTTAGATTGAATTT 1289
Db 146300 -----ATGTTTGTGCT 146311
QY 1290 TTTCTCCTAGCCTTTTGGGTCGATTTGAAGATAGATATTTCTTTTACATCTGATTTTA 1349
Db 146312 TCCTTCAGGAGCTCTTTTAGGTCAGCCCTGGTGTGTGACAAATCTCTCAGCATTTGCTTG 146371
QY 1350 TCTTAGATGCTCTTTCTCTCAACTATTTGTGACAGAAAGTTTCTTAAGTGCAGTAGT 1409
Db 146372 TCTGTAAGTATTTTATTTCT 146431
QY 1410 CTGSCCTGACATCTGTAGTCTCTGTGAGTCTGTAGCACATCTGTGAGGCTCTCTTACA 1469

Db 146432 CTGGGTTGAAATTTCTTTTCTTTTAAAGATGTTGAATTTGGCCCCCACCCTCTCTCTGCT 146491
QY 1470 TTTTGTAGTTCTATTTGGAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1529
Db 146492 TGTAGAGTTTCTGCT--GAGAGATCCGCTGTTAGTCTGATGGGCTTCCCTTTGT--GGGTAA 146549
QY 1530 TTTGCTTTTCTTTTCCCTTTGCACTTTTAAATTTCTTTTCTTTTCTTCTATATCTTTTAGTATT 1589
Db 146550 CCCAACCTTTCTCTGCTGGCTGCCCTTAAACATTTTCTTCTTCACTTTTGGTGAATC 146609
QY 1590 TGAATTTATGACACTGTGGGAGTCTTTTTCGGGTCCAACTATTTTGTGTTTGTATG 1649
Db 146610 TGAATAATTAATGCTGTTGAGTTGTTTCTCGAGAGTATCTTTGTTGTTCTCTCTGA 146669
QY 1650 CTCTTGTACCTTGTATGAGCACTCTCTTCTCAAGGTTAGGAAATTTTCTTTTGTGTTT 1709
Db 146670 TTTTCTGAATCTGAATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146729
QY 1710 TCTGAAATATTTTCC-----CTGCTTTTACCTGCTCTTTTCCCTTCC 1755
Db 146730 CTTGGGAGTGTTTTCCAACTTTGGTTCCATTTCTCCGCTCACTTTTCAAGTACACCAATCG 146789
QY 1756 TCTATTTCTTTTCTTTTCTGATGCTCTGCTGCTTCTGCTGCTTCTGCTGCTGCTGCTG 1815
Db 146790 GACGTAGATTTGTTCTTTTCCATAGTCCCATTTCTTTGGAGGCTTTGTTGTTCTTT 146849
QY 1816 TTTTGTAGCTTAACTATTTCTTTTACCAAGTATCCATTTCTTCTATCTTCTTCTTCACTGC 1875
Db 146850 TTATTTCTTTTCTCTAAATTTTCTTCTCACTTTCACTTTCACTTTCACTTTCACTTTCA 146909
QY 1876 CTGAGATCTCTCTTCTATCTCTGTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1921
Db 146910 CTGATAACCTTTCTTTCCAGTTGATCATCATCTGCTGCTGCTGCTGCTGCTGCTGCT 146955

RESULT 20
US-09-816-095-3/c
; Sequence 3, Application US/09816095
; Patent No. 6664084
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) .. (99916)
; OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

Query Match 9.0%; Score 204.799; DB 1; Length 99916;
Best Local Similarity 50.1%; Pred. No. 5.5;
Matches 831; Conservative 0; Mismatches 722; Indels 106; Gaps 14;
QY 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGTGTTGTTTCC 351
Db 40959 TTTATGCTTCTATTTGCTTATTTACCCAGTAGTCACTTCCAGAGAGGTTGTTCACTTTCC 40900
QY 352 ATAAGTTTGTAAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 40899 ATGAGTTTGTCCAGTTTGTGAGTGAGT---TTCTTAATCCTGAGTCTTAATTTGATTGCAAT 40843
QY 412 GGTGTCAGATGAGACATAGAGTATTAATTTCAATTTGCTTTTATCTGTCGAGACTTCTCTT 471

Db 40842 TGGGGTCTGAGAGACAGTTTGATTAATTTCTGTTCTTTTATATATTTGCTGAGAGTGT 40783
QY 472 TGTGTTGAAATATGATTAATCAATTTTGGG--GAGTTTTCATAGGCTGCTGACAAAGAGTAC 529
Db 40782 TGTCTCCAAATATGATGCTCAATTTTGAATAAGTGTGATGGGCTGAGGAGATGTAT 40723
QY 530 AGTCCTTGTGTTTGTGTAAGTAG--TCTGTAAATATCTCTAGTCCACTTGGTTATGAC 588
Db 40722 ATTCTGTTGATTTGGGAAGAAGAGTCTGTAGATGCTATTAGTCTTGTCTGTCTAGAG 40663
QY 589 ATCAGTTAGTCTCAGCATTTCTCTGTTTCTGTTTGTGTTGATGACCACTAACTGTTGGAG 648
Db 40662 CTGAGTTCAAGCTCTGGATATCTTTGTTAACTTTCTGTTCTGTTGATCCATCTAATATTG 40603
QY 649 AGNATGGGTATTTGAAGTAGCCCACTATCTGTGTGTAG--GTCAATATGATTTTAGCT 707
Db 40602 ACAGTGGGTGTAAAGTCTCCCATATATATTGTGTGGAGTCTAAGTCTCTTTGTAGAC 40543
QY 708 GTAGC-----TGTGCTTGTGTTTATGAACCTTGGTGACATTTGTTGGTGATAGACATT 762
Db 40542 TTAGTCTCTAAGAACTTGTCTTTATGAATCTGGTGCTCTCTGTTTGGTGATATATT 40483
QY 763 AAGAAATGCAATGCTCTTGTGTGATT--TTCTTTGATGCCCTATGATGATTTCTCCC 820
Db 40482 TAGGATAGTTAGTCTCTTGTGTTGAATGATCCCTTTTACCAATATGTTAAGGGCTTTCTT 40423
QY 821 AATCTCATCTGTTAGTTTGGGTTTAAAGTCTATTAGTCAGATATTAATAATGACTGTATC 880
Db 40422 TGTCTCTTTGATCTTTGTTGTTTAAAGTCTGTT--TCATCAGAGACAGGATGGCAA 40365
QY 881 GGCCTGCTCTTAGGGCCATTTGCTTAGAATAATCTTTTCCATPCCTTTTACTCTAAGGTGA 940
Db 40364 CCGTTTCTTTTCTTCCATTTGCTTGTGTAGATCTTCTTCCATCATTATTTTGGAGCTA 40305
QY 941 TGTCTATCCATCGTAGGTGCTCTTTTGGATGACAGTAGGATGATCTTGTGTTTCAT 1000
Db 40304 TGTGTCTCTTGTGATGATGATGATGCTCTCTGATACAGTC--AAGGCTCTGACTCTT 40247
QY 1001 ATCCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGATGATGATGT 1060
Db 40246 ATCCAAATTTGCCAGTCTGTGCTTTTAAATGGGGCAITTTAGCCCACTTA----- 40198
QY 1061 TGAGAATTATCAATGACAGGTGTTTGTGGAATCTTGTGTTATCTGCACTTGTGAAGTGT 1120
Db 40197 -----CATGAGGTTAATATTTGTTGTTGA 40170
QY 1121 GT 1180
Db 40169 ATTTGATCTGTGATATG-----ATGTTAGCTGGTTATT 40135
QY 1181 GT 1240
Db 40134 TTGCCCAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 40075
QY 1241 TTATTTATCAATTTCTTGAATGTTGGGTAAATCTTTTGTAGATGAAATTTTCTCTAGC 1300
Db 40074 TTTGTAGTGGCTGT 40020
QY 1301 CTTCCTTAGGCTGATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTTAGATGT 1360
Db 40019 CTTTGTAAAGCAGGCTGT 39960
QY 1361 CTTTCTTTCTCCAACTATTTGTGACAGAAAGTTTCTTAAAGTGAAGTGTGTGTGTGTGT 1420
Db 39959 TTTTATTTCTCTTCACTTATGAAGCTTAGTTTGGCTGGATATGAATTTCTGGTTTAAA 39900
QY 1421 TCTGTAGTCTCTGTGAGTCTGTAGACATCTGTGACAGGCTCTTCTTACATTTTGTAGTTTC 1480
Db 39899 ATTCTTTCTTTTAAAGCATTTGAATATCGGCCCTCTCTCTCTGCTGTGTGTGTGTGT 39840
QY 1481 TATTGGAAAGTCAAGTGTATTTCTTAATATCATCTGCTTTTATATGTTAATGCTTTT 1540
Db 39839 TGCT--GAGAGATCTGCTGTTAGTGTGAGTGGCTTTCCCTTTGT--GGGTAAACCCGACCTTTC 39782

QY 1541 TCCCTTGATCTTTTAAATATTTCTTTTGTGTTCTATATCTTTTGTGATTTGATTTATG 1600
Db 39781 TCTCTGGCTGCCITTAACATTTTCTTCTTCACTTTCAACCTTGGTGAATCTGACAAATATG 39722
QY 1601 CACTGTGGGAGTTCCTTTTCCGGTCCAAATATTTTGGTGTGTTGATGCTTCTTGTACC 1660
Db 39721 TGTCTGGGGTGTCTTCTCTGAGGAGTATCTTTGTTGGTGTCTCTTCTTCTGAAAT 39662
QY 1661 TTGATAGGCATCTCTTTCTCAAGGTTAGGAAATTTTCTTTTGGTGTCTTCTTCAAAATA 1720
Db 39661 TGAATTTGGGCTGCCITTTGCTAGGTTGGGAAGTTCTCTCTGATATATCTCTGAAAGTGT 39602
QY 1721 TTTTCCCTGCTTTTGACCTGCTTCTTCCCTTCTCTATATTC-----TTTG 1767
Db 39601 TTATPAACCTGTTTCCATTTCTCCCTGTCACTTTCCAGTACACCAATCAAAACGTAGATTG 39542
QY 1768 GTTTTTCATAGTGTCTCTGCTTCTCTGGATGTTTATGCTGATTTATTTAGACTTAA 1827
Db 39541 GTCITTTACATAGTCCGATTTCTTTGGAGGCTTTATGCTTCTTTCACCTCTTTT 39482
QY 1828 CATTTTCTTTGACCAAGTATCCATTTCTTCTATCTTGTGCTTCTGCTGCTGAGATCTCT 1887
Db 39481 CTCTAATCTTCTTCTTGTCTTTATTTTCAATTTGATCTTCAATCACTGATATCTTT 39422
QY 1888 CTCTATCTCTGTTATTTCTGTCAGTGAAGGCTTGTCTCTG 1926
Db 39421 CTTTCGCTTGATTCATCAGCTATTGAGCTTGTGTATG 39383

RESULT 21
US-09-801-876B-3/c
; Sequence 3, Application US/09801876B
; Patent No. 6492155
; GENERAL INFORMATION:
; APPLICANT: YE, Jace et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CLO01160
; CURRENT APPLICATION NUMBER: US/09/801.876B
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) -- (148567)
; OTHER INFORMATION: n = A, T, C or G
US-09-801-876B-3

Query Match 8.8%; Score 200.5; DB 1; Length 148567;
Best Local Similarity 50.8%; Pred. No. 4;
Matches 894; Conservative 0; Mismatches 710; Indels 155; Gaps 19;
QY 305 TTCTATCTTGGCTCATTTTAACTCAGTAGTGTGTTTGGTTTCCATTAAGTTTGTAG 364
Db 45036 TTTTGTGTTTGAACCAATGCCATTGAGGACAGGTTATTATTTCCATGTTTGTGATG 44977
QY 365 TTTTCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 424
Db 44976 GTTT---TGAAGTTTCTTTTGGAGTTTCAATTTTATTTCCACTTTTATTTCCACTGTGCTGAGAG 44920
QY 425 GACATAGAGTATTATTCAATTTCTTTTATCTGTCGAGACTTGTCTTCTTTTGAATAAT 484
Db 44919 AATGTTGA--TATTAATGAATATTTCTAATTTATTGAGCTCAATTATGAGCTTATCAT 44861
QY 485 GTATTTCAATTTGAGAG--TTTTCATAGGTTGCTGACAAAGAGTACAGTCTTT--GTGTT 541
Db 44860 ATGCTCTATCTTGAGAAACCTTCCATGCTCTGTTGTAATGAATGTGTAATCTGCACTTCT 44801

485 GTATTCAATTTTGGAGAG--TTTCATAGGCTGTGACAAAGGTACAGTCTTT-CTGTGT 541
4860 ATGCTCTATCTTGGAGAACTTCCATGCTCTGTGTAATAGATGTATCTGCAGTTGT 44801
542 TTGCTGAAATAGCTGTAAATATCT-CTAGGTCCACTTGGTTTATGACATCAGTTAGCTC 600
44800 TCGATGAAATCTTCTGTATATCTGTATAATCCATTTGT-TTCAAGGTATAGTTAAATC 44741
601 CAGCATTTCTCTGTTTCTGTTTGTGAGATGACCTAACTGTGAGAGAAATGGGGTAT 660
44740 CATTTGTTCTTTG-TTGACTTTCTCTGATGACCTGTCTAGTGTCTGAGTAT 44682
661 TGAAGTAGCCACATCTCTGTG-TGAGGTCAATATGATGATTTAGCTGTAGCTGTGCTT 719
44681 TAAATCTCTCACATATTATTGTTGTTGCTGTCTCTCATTTCTTAGGTCAATTAGTAAT 44622
720 GTTTTATGAACCTTGGGTGACATTTGTTTGGTGCATAGACATTAAGAAATGCAATGTCCT 779
44621 GTTTTATAAATTTGGGAGCTCCAGTGTAGGTGCTTATATGTTTATGATTTGTGACATTT 44562
780 CTGCTGTGA-TTTTCTTTGATGCTATGTAGTATTTCTTCCCATCTC---ATCTGCTTA 835
44561 CCGTTGNAACAAGCCCTTTACCATATATAATCTCCCTCTTTGCTCTCTTTAACTGC 44502
836 GTTTTGGGTTTAACTATTAAGTCAATATTAAGTCAATATTAAGTCAATATTAAGTCAAT 895
44501 TGTACTTTTAAAGTTTGTGTTGCTGATATAAGAAATAGCTACCCCTCTCTGCTTTGGT 44442
896 GCCATTTGCTTATAGATATCTTT-TCCATPCTTTTACTTAAAGTGATGCTATCCATGG- 953
44441 TCCATTTGCAATGAAATTCCTTTCCCAACCTTTTACTTTATGTTTATGTTAGTACATG 44382
954 --TAGGTTGCTTTTGGATGCAAGCAGTAGATGGAATCTGTTTGTGTTTCAATPCCATCTGT 1011
44381 TACTAGGTGAGTCTCTGAAGCAGCACATAGTAGGTGTTGAGTCTTATPCCATCTGT 44322
1012 TACCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCAATGATTTGAGAAATATC 1071
44321 GATCTGTATCTTTTAAAGTGAAGATTTAGGCCATTTTACATCCATGTTAGTATGAA 44262
1072 AATGACAGTGTGTTGATCTTTGTT----- 1098
44261 TGTGAGGTACCGTTGGCATTCATCATGCTCTTTGTTGCCCTGTGTTGTTGGTCTTTT 44202
1099 -----ATCTTGACATTTGAA 1114
44201 CCAAAAAACCAAAAGAGCTAAAGCTGTTCTTCCCAAAAGACCTTCAGCTTCTCCA 44142
1115 GTGCTGTGTGTGTGTG-----TGTTGTGTGTGTG 1145
44141 GTGGGGTGTGTGTGTGAGGAAGATGCTCTTCTTCTTCTTCCACAGTTGGAGCACT 44082
1146 TGT 1205
44081 CACATTTTGTGGGGT 44022
1206 TCTCTCCCTCTTTTGTATTTTGGCT-----GGAATTTTATTTATTTATTTTCTT 1259
44021 TTTAATCTGTATTTTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 43962
1260 GAATGTGGTAACATCTTTAGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1311
43961 TTGATGTGTTTCCAGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 43902
1312 CTGCAATTTGAAGATAGATATCTTTTACATCTGATTTTATCTAGATCTCTTTCTTCTC 1371
43901 GTGACTTGTATGGAATTTTCTCTGAGATTTGTTGTGTGAAATGACTGTATCTTTC 43842
1372 CAACTATTCTGACAGAAATTTTCTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1431
43841 CTTTATATATGATCTTATCTTCACTGGGATACAAAATCTTGGCTAAATATTTTGTAG 43782
1432 TTGGAGTGTGTAGACATCTGTGTGAGGGCTTCTTACATTTTGTAGTTTCTATTGAAAG 1491

43781 GAGGCCAGAG----ATAGGGTCCAATCCCTTTAGCTTTAGGTTTCTGCTAGAAA- 43727
1492 TCAGGTGTAAATCTAAATACATCTGCTTTATATATGTAATTTGTTTCTTTTCCCTTGGATC 1551
43726 TCTGCTGTAAATCTAAATAGGTTTCCCTTTGAGGTTTCTGCT- GCTTCTGCTCATATC 43668
1552 TTTTAAATCTTCTTCTTCTTCTATCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1611
43667 TCTTAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 43608
1612 GTTCTTCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1671
43607 ATACCTTTTGGGATGAATTAACCAAGGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 43548
1672 CTCCTTCTCAAGGTAGGAAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1731
43547 ATTTCTAGCAAGCGCGGAAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 43488
1732 TTTGACCTGCTTCTT-----CCCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1780
43487 TTTAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 43428
1781 GTCTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1840
43427 ATATCAGACTTCTTGGAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 43368
1841 CAAGGTATCCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1900
43367 TCAATGGGTAAATTTGAAGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 43308
1901 TATTTCTGCTGAGGCTT 1919
43307 AATTTCTGCTGACATTT 43289

RESULT 23
US-09-426-290-1/c
; Sequence 1, Application US/09426290
; Patent No. 6410712
; GENERAL INFORMATION:
; APPLICANT: Berglind Ran Olafsdottir
; APPLICANT: Jeffrey Gulcher
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
; FILE REFERENCE: 2345.2001-000
; CURRENT APPLICATION NUMBER: US/09/426,290
; CURRENT FILING DATE: 1999-10-25
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 168575
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (21181)...(21403)
; NAME/KEY: CDS
; LOCATION: (95252)...(95430)
; NAME/KEY: CDS
; LOCATION: (101753)...(101996)
; NAME/KEY: CDS
; LOCATION: (110324)...(110439)
; NAME/KEY: CDS
; LOCATION: (124058)...(124278)
; NAME/KEY: CDS
; LOCATION: (127009)...(127130)
; NAME/KEY: CDS
; LOCATION: (128910)...(129139)
US-09-426-290-1

Query Match 8.4%; Score 191.799; DB 1; Length 168575;
Best Local Similarity 49.2%; Pred. No. 3.8;
Matches 816; Conservative 0; Mismatches 737; Indels 104; Gaps 13;


```

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001067
; CURRENT APPLICATION NUMBER: US/09/751,389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

Query Match          7.7%; Score 176; DB 1; Length 786431;
Best Local Similarity 50.4%; Pred. No. 1;
Matches 889; Conservative 0; Mismatches 755; Indels 120; Gaps 20;

QY 292 TCTATTCTTGATTTCTATCTTGCTCATTTTAACTCAGTAGTAGTCTTGTGGTTTC 351
Db 653422 TTCTGCTTAAAGTTCAATTATACCAAAAGTCATTCCAGGAGGATATACAAATTTCC 653363

QY 352 ATAAGTTTGAAGTTTCTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 653362 ATGTAATGTGTGTTTGTAGTGAT---TTCTTAGTCTTGGATTCGAAATTTGATGTC 653306

QY 412 GTCGTGAGATAGACATAGATGATATTTCAATTTGTTTATCTCTCAGACTTGCCT 471
Db 653305 TGCGCTGAGAGATGT----TAAGATATGATTTCTTTTACATTTGCTCAGGAGTAT 653250

QY 472 TGTTTGAATATGATTTCAATTTTGAG--AGTTTCATAGGGTGTGACAAAGAGTAC 529
Db 653249 TACTTCCAATATGATCAATTTTAGAGTAAGTGCCATGTGTGACGAGAAATATGAT 653190

QY 530 AGTCT-TTGTGTTTGTGAATAGTCTGTAATATCTCTAGTGCCACTTGGTTTATGAC 598
Db 653189 ATGCTGTGCATGGAGTGGAGAGGCTTTAGATATCTATCAAGTCCATTCATCCAGAG 653130

QY 589 ATCAGTTAGCTCCAGCATTTCTCTGTTTGTGTTTGTGAGATGACCTAACTGTTGGAG 648
Db 653129 CTGAGTTCAGGCTCTGAATATCTTTGTGAATTTTCTATCTGATGATCTGCTAAATTC 653070

QY 649 AGAATGGGATTTGAAGTAGCCCTATCTGTGTGTGAG-GTCAATATGATTTTAGCT 707
Db 653069 ACAGTGGGGTTTAAAGCTTTCACATATTATGTGTGGAGTCTAAGTCTCTTTCAAGGT 653010

QY 708 GTAGCTGTGCTGTTTATGAACTTGGGTGACATTTGTTTGGTGCATAGACATTAAGAA 767
Db 653009 CTCTGATAACTTGTCTTATGAACTGAGCGCTCTGTG-TTGTGTCATATATTTAGGA 652951

QY 768 TTGCAATGCTCTCTGTGAGTAT-TTCTTTGATGCTCTATGATGATTTCTTCCCAATCTC 826
Db 652950 TAGCCAGCTCATCTTGTGTAATGAAACCCCTTACCATTATGTAATACCTTTCTTTGTCTT 652891

QY 827 ATCTGCTTAGTTTGGGTTTAAAGTCTATTAGTCAGATATTAATGACTGTATCGGCTTG 886
Db 652890 TTTTGAATCTTTGTTTAAAGTCTTTTGTGCAAAACCTGGATGCGCAACCTGTTTT 652831

QY 887 CTCTTAGGGCCATTTGCTTGAATATCTT-TTCCATCTTTTACTCTAAGGTGATGCT 945
Db 652830 TTTCTGTGCTCAATTTGCTTGTAAATTTTCTGTCATCACTTTATTTTGGCCCTACCTGT 652771

QY 946 ATCCATG-----GTAGTTGCTTTTTCGATGCGAGTAGATGATCTTTGTTTTTC 998
Db 652770 GTTACTGATGTAGTGAGTGGGTCTCTTGAAGACAGATATTGATAGGTCTTAGTTCT 652711

QY 999 ATATCCATTTCTGTTACCAGTATCTTTTCTTAGAGAAATTAAGATCATTTAGTCAATTCAT 1058
Db 652710 TTAATCCAGCTTG-----CCATTTTATCTTTTAAATGAGGCACTTAGCCCAATTTAC 652658

```

RESULT 29
US-09-497-855A-40/c
; Sequence 40, Application US/09497855A

```

QY 1059 GTTGAGAATTAATCAATGACGAGTGTGTTGTTGATTTCTTGTATCTTGCACCTTGTGAAGTGT 1118
Db 652657 ATTTTCAGGTTAGTATTGTTATGTTGGATTGATTCCTGACATCGTATGCTAGTAATTA 652598

QY 1119 GTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1178
Db 652597 TTTTGCAGACTTGTTTATGCTGTTAATTTATAGTGTCACTGT----- 652556

QY 1179 CTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1238
Db 652555 -----TCTGTGTACTTCAATTTGTTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 652516

QY 1239 ATTTTATTTATCATATTTTCTTGAATGTGGGTAAACATCTTTAGATTTGAAGTTTCTCTCTA 1298
Db 652515 GTTTTT-----CCTTTCTGCAITTAGGCTTCTCTCAGG 652483

QY 1299 GCCTTCTTTAGGTCGTGCAATTTGAAGATAGATATTTCTTTACATCTGATTTTATCTTAGAAT 1358
Db 652482 AGCTCTTATAGACAGGTCGTGTGTGATGAATTCCTCAGCAITTTACTTGTCTGAAAAG 652423

QY 1359 GTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAAAGTCAGTAGTCTGGCCTGA 1418
Db 652422 GACCTTATTTTCCCTTTCTTATGAAGCTTAGTTTGAATGGATATGAAATTTCTGGGTGG 652364

QY 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGACGGGCTTCTTACATTTTGTAGTT 1478
Db 652363 AATCTCTTCTTTTAAACAATGTTGA-----ATGCCCCCAATGCTTGTGGCTTATAAGGTT 652308

QY 1479 TCTATTGAAAAGTCAGGTGAATTTCTAATAATCATCTGCTTTTATATGTTAATTTGTTCTTT 1538
Db 652307 TCTGCT-AAAAGGTTCTCTGTAGTCTGATGGGCTCCCTTTTGTAGGTGATCTGG-CCCT 652250

QY 1539 TTTCCCTTGCATCTTTTAAATATCTTCTTGTCTTATCTTGTATCTTGTGATTTGATTTGATTA 1598
Db 652249 TCTCTCTGACTGCCATTAACATTTTCTCTGTATTTGACCTTGGAGAACTGATGATTA 652190

QY 1599 TGCACTGTGGGAGTCTCTTTTCCGGTCCAATCTTATTTGGTGTGTTTGTGATCTTCTTCTGTA 1658
Db 652189 TGTGCTCTGAGGATAATCTTCTCAITTAGATATTTTACTGGGGTCTCTCTGCATTTCTCTGAA 652130

QY 1659 CTTGATAGGCACTCTTCTTCAAGTTAGGAAATTTTCTTTTGTGTTTCTTGTGTTTCTTGA 1718
Db 652129 TTTGAATGTGGCCTGTCTAGCTTGGTGGGTAAGTTCTCATGATGATATTTCTGAAATA 652070

QY 1719 TATTTTTC-----CCTGCTTTTGACCTGCTCTTCTTCCCTTCTCTTCTTCTTCTTCTTCT 1764
Db 652069 TGTTTTCTACATTTGTTCCATTTCTCCCATCTCTTTTACAGTACCAATCAGTTGTAGAT 652010

QY 1765 TTGGTTTGTGATAGTCTCTGCGCTTCTGATGATTTTATGCTGATTTTATTTAGACT 1824
Db 652009 TTGAATCTTTTACATAAATCTTATTTCTTGGAGGTTTGTATTCTATTTCTCAITTTTTT 651950

QY 1825 TAAACATTTCTTTGACCAAGGATCCATTTCTTCTATCTTGTCTTCACTGCTGCTGATTC 1884
Db 651949 TTCTCTATTC-TTGTCTACCAGTGTATTTTCAAGAACCCAACTTTAGTCTGATTC 651891

QY 1885 TCTCTTATCTCTTGTATCTGTCTGAGTGGCTTGTCTCTGAGGTTCTCTGTGGGTTCT 1944
Db 651890 TTTCCCTATTTGGTCTATTTCTGTATTAAATCTTGTGATTTGATTAATAAAT---TCTT 651834

QY 1945 TAAATTTTCTTCTTCAAGTATCTTCTGATTTGCTTGGGTTTGTGTTTATTAAT---CTATTTCC 2002
Db 651833 TTAATGTGTTTTTGTAGTCTGCAATCAGTTCAAGTTAGTTGTTCTCTATGCTAGCTATTTTG 651774

QY 2003 ACTTTTCAGGCTCTGAAATGTTTTTA 2026
Db 651773 ACTGTGAGCTCTGCAATGTTTTTA 651750

```


Sequence 3, Application US/09851896
Patent No. 6410325
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Susan M. Freier
APPLICANT: Andrew T. Watt
TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT)
TITLE OF INVENTION: EXPRESSION
FILE REFERENCE: RTS-0220
CURRENT APPLICATION NUMBER: US/09/851.896
CURRENT FILING DATE: 2001-05-08
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 3
LENGTH: 70000
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
US-09-851-896-3

Query Match 6.6%; Score 149.3; DB 1; Length 70000;
Best Local Similarity 56.0%; Pred. No. 14;
Matches 431; Conservative 0; Mismatches 322; Indels 17; Gaps 8;

452 TTATCTGTGACAGCTGCTGTTTGAATAATATGATTTCAATTTGGAGA--GTTTCATA 509
Db 24271 TAATAATTAGACTGTTGTTGGCCCTACCATATGCTCTATCTTGAGATGATCCATA 24330

510 GGGTGTGACAAAGGTACAGTCT-TTGCTGTTTGGTGAAATAGTCTGTAATATCTCT 568
Db 24331 TGCTGAGGAGAAGATGTTATCTGCAGCCGTTGGATGAATGTTCTGTAAATATCTAA 24390

569 AGGTCCACTGTTTATGACATCAGTTAGTCCAGCATTTCTGTTTCGTTTTTTGTTG 628
Db 24391 TAGGTTCACTGATCTATCTGCAGACTAAGCGTGGTTCTTTGTTGATTTCTGCT 24450

629 AGATGACCTAACTGTGGAGAGAAATGGGTTATGAAGTAGCCCACTATCTGTGT-GTGAG 687
Db 24451 GGATGAACCTGTCTAATGCTGAAGTGGAGTAGTGAAGTCTCCAGCTATTATGTTAG 24510

688 GTCAATATGTAATTTAGCTGACTGCTGTTGTTTATGACTTGGTGGTGCACATTTGTT 747
Db 24511 GTCTATCTCTCTTTAGCTCTAAGAACTATCTCTTATATATCTGGGTGTTCCAGTGT 24570

748 TGGTGATAGACATTAAGAAATGCAATGCTCTTGTGGTGAATTTTCCTTTGATGCCCTATG 807
Db 24571 GAGTGCAATGTAATTAAGAAATGTTACATTAATCTTGTGTAATTCACCTTT- - -TTATA 24625

808 TAGTATTTCTCCCAATCTCATCTGCTTAGTTTGGGTTTAAAGTCTATTT-AGTCAGATA 864
Db 24626 TAATGACCTCTCTTGTCTCTTTTATAGTTTGTCTTGAAGTCTGTTTGGCCCTCATAGA 24685

865 TTAAATGACATGATCGGCTTCTTCTTTAGGCCAATTTGCTTGAATAATCTTT-TCCATC 923
Db 24686 ATAGCATAGCTATTCCTGCTCTTATTTGGTTTCCATTTGCAATGCAATATCTTCTCCATC 24745

924 CTTTACTCTAAGTGATGCTATACCAAT- - -GGTAGGTGCTGTTTGGAGTGCAGGAGT 980
Db 24746 CTTTATGTTTTCAGTCTATGCTGCTTTTACAGGTAAGATGTTGTTTCTATGAGGCAACAGA 24805

981 AGGATGGAATCTGTTTTCATATCATCTGTTTACCCAGTATCTTTTCTAGAGAAATTA 1040
Db 24806 TCAGTGGGCTGTGTTTCTATCCGTTCTGCCATCTGTGTTGATTTGATTTGGAGAGTTTA 24865

1041 GATCATTTGAGTCAATGATTTGAGAAATTAATCAATGACAGTGTGTTGGATTTCTGTTAT 1100
Db 24866 GTCCATTTA-TGTTCAATGTTATTACTGACAAAGTAAAGACTTACTCTCCGCACTTTATTA 24924

1101 CTGCACTGTGAGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1160
Db 24925 ATTAATTTCTGTGTTGTGATCTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 1160

1161 GT 1210

Db 24985 GTGATTTTCTCTGTTGGTATGTTTAAATTTCTGCTTTTATTTTGTGT 25034

RESULT 31
US-09-803-671B-3/c
Sequence 3, Application US/09803671B
Patent No. 6582946
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001161
CURRENT APPLICATION NUMBER: US/09/803.671B
CURRENT FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 64467
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(64467)
OTHER INFORMATION: n = A,T,C or G
US-09-803-671B-3

Query Match 6.4%; Score 145.4; DB 1; Length 64467;
Best Local Similarity 50.9%; Pred. No. 16;
Matches 595; Conservative 0; Mismatches 531; Indels 42; Gaps 11;

292 TCTATTTCTTGTATTTCTATCTATCTTGGCTCAATTTTAACTCAGTAGTGTGTTGTTTCC 351
Db 54901 TTTTCTTCTTAAATTTCTTCTTGGCCAGTAGCCATTCCAGCAATGTTGTTTGTTC 54942

352 ATAAAGTTTGTAAATTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 54841 ATGATTTATACAGTTTCCAAAGTTCCTCTGTTATGATTTCTAGTTTATTTCCATGTT 54782

412 GGTGCTCAGATAGGACATAGATATTTTCAATTTGCTTTTATCTGTCGAGACTTGCTT 471
Db 54781 GG- - -CTGAAGAATATCTGTAATGATTTTGAATTTTAAAAATTTTGAACCTGTTT 54726

472 TCTTTGAAATATGATTAATCAATTTTGGAGAGTTTCAATAGGTTGCTGACAAAGGTACAG 531
Db 54725 TGCAGCCTAAATGATATGGCTATCTCTGAAATGTTTCCATGTTTAAAGGAACTTATTT 54666

532 TCTTTGTTGTTTGGTGAATAGTCTGTAATATCTCTAGTCCACTTGGTTTATGACATC 591
Db 54665 CTGCCATTTGTGAGAAATATCTCTAGGTATCTGTTGTCTGTTGTTGTTTAAAGTGC 54606

592 AGTTAGTCTCCAGCAATTTCTGTTTTCGTTTGTGTTGAGTACCTAATCTGTTGAGAGA 651
Db 54605 AATTTAAATCCAGGTTTCTCT-ATTGATGTTTGTGTTAGATGATCTGTTTAAATGTTGAGA 54547

652 ATGGGTTATTAAGTAGCCCACTATCTGTTGTTGAG-GTCAATATGATTTTACCTCTA 710
Db 54546 GTGGGTTGTTAAAGTCCCATCTATGATTTATAGATTTATCTCTCTCTTCAGATCTA 54487

711 GCTGTGCTTTGTTTATGAACCTTGGGTGACATTTGTTTGGTGCATPAGACATTAAGAATG 770
Db 54486 ATAAATTTGCTTATATACCTGGGTGCTGCTGTCAGGGTGCATC- - - - -AGAACTG 54434

771 CAATGTCTCTTGGTGGATTTTCTCT-TTGATGCCCTATGATGATTTCTTCCCAATCTCATC 829
Db 54433 TTACATCTCTGCTGCAATTTTCTCTTTATCATTTATATAGATTTGTTCTCTTTTAC 54374

830 TGTGTTAGTTTGGGTTTAAAGTCTATTAAGTCAATTAAGTCAATTAAGTCAATTAAGTCA 889
Db 54373 TGT- - -TTTGACATAAAGTGTGTTTATCTGATATGAATATAGTACTCATGCTCTCTA 54317

890 CTTAGGGCCATTTGCTTGAATATC-TTTTCCATCTTTTACTCTAAGGTGATGCTATC 948

```

RESULT 32
US-09-393-634-8
; Sequence 8, Application US/09393634
; Patent No. 6558910
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: SF. a No. 6558910el Family of Taste Receptors
; FILE REFERENCE: 02307E-098000US
; CURRENT APPLICATION NUMBER: US/09/393,634
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1540
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat GR04 sequence approximately 1100 bp 5' to SEQ
; OTHER INFORMATION: ID NO:92
US-09-393-634-8

```

Query Match 6.3%; Score 142.8; DB 1; Length 1540;
Best Local Similarity 67.1%; Pred. No. 6.3e+02;
Matches 286; Conservative 0; Mismatches 132; Indels 8; Gaps 6;
QY 589 ATCAGTTAGCTCCAGCATTTCTCGTTTCGTTTTTGGTTCAGATGACCTAACCTGTTGGAG 648

```

RESULT 33
US-09-220-132-168/c
; Sequence 168, Application US/09220132
; Patent No. 6506607
; GENERAL INFORMATION:
; APPLICANT: ShyJan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 168
; LENGTH: 9573
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-168

```

	Query Match	6.0%;	Score 136.8;	DB.1;	Length 9573;
	Best Local Similarity	57.6%;	Pred. No. 1.1e+02;		
	Matches 333;	Conservative 0;	Mismatches 237;	Indels 8;	Gaps 5;
QY	292	TCTATTCTTGATTTCTATCTCTGGGTCAATTTTAACTCAGTAGTCAGTTGTTGGTTTCC	351		
Db	4147	TTTCCTTCTTGATTTTCATTTGTTAAACCAGGTCATTCAGGAGCAGATTTATTAATTTTC	4088		
QY	352	ATAAGTTTCTAAGTTTTCGTGTTCTCGTCTGTTGTTGTTATCTAGATTTAAGCTGT	411		
Db	4087	ATGTAATTCGTATAGTTTTAAAGGT--TCTTTTGGAGTTAATTTCCAGTTTATTCAT	4031		
QY	412	GGTGTCAGATAGGACATAGAGTATATTTCAATTTGCTCTTTTATCTGTCGACACTGCTT	471		
Db	4030	TGTTGGTCTCGAAGAGGTACTTTCACATTAATTTCAATTTTCTTAATTTTGTTCGACACTGTTT	3971		
QY	472	TGTTTTGAATATGTATTCAATTTTTCGAGA--GTTTTCATAGGGTGTGCACAAGAGGTAC	529		

Db 3970 TGTGGCCTATCTTATGATTTATCTTGGAGAACGTTCTATGCTGCTGATTAATAATGGTAT 3911
Qy 530 AGTCT-TTGCTGTTTGGTGAATAGTCTGAATAATCTCTAGTCCACATTTGGTTATGAC 588
Db 3910 ATTCGAAAGTTTGGGTAGAGTTAGGTAAATAATCTGCTAAATCTATTTGTTCTAGGA 3851
Qy 589 ATCAGTTAGCTCCAGCAATTTCTCTGTTTCGTTTTTGTTCAGATGACCTAACCTGTTGGAG 648
Db 3950 TATTAATTAAGTCAATTTTCTTTATTTGACTTTCTGCTTTGCTGCTGCTAGTCTG 3791
Qy 649 AGAATGGGATTAAGTAGCCCACTATCTGTGT-GTGAGGTCAATATGATGATTTTAGCT 707
Db 3790 TCCATGGAGTATGAAGTCCCTACTATTAATTTGTGGTGCATCTATCCCAATTTCTTAGGT 3731
Qy 708 GTAGCTGCTGTTTATGAACCTTGGTGACATTTGTTGGTGCAATGACATTAAGAA 767
Db 3730 CTAGTAATTAATTTTATTAATTTGGAGCTCCAGTATTAAGTGCAATATTAATTTAGGA 3671
Qy 768 TTGAATGTCTCTTGTGGGA-TTTTCTTTGATGCTATGATGATTTCTTCCCAATCTC 826
Db 3670 TTGTGATATTTGTTCTGTTGGACCATCTTTATCATTAATAAATTTCCCTCTTTGCTT 3611
Qy 827 ATCTGCTTAGTTTGGGTTTAAGTCTATTAGTCAGATA 864
Db 3610 TTATAACTGTTGTTGCTTTAAAGCTTGTCTGATATAAA 3573

RESULT 34
US-09-146-053-4
; Sequence 4, Application US/09146053A
; Patent No. 6399349
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; APPLICANT: Sprinkle, Terry Joe Curtis
; APPLICANT: Venema, Richard C.
; TITLE OF INVENTION: Human Aminopeptidase P Gene
; FILE REFERENCE: MGI103
; CURRENT APPLICATION NUMBER: US/09/146,053A
; EARLIER FILING DATE: 1998-09-02
; EARLIER FILING DATE: 60/057,854
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-053-4

Query Match 5.5%; Score 125.5; DB 1; Length 50000;
Best Local Similarity 54.2%; Pred. No. 25;
Matches 429; Conservative 0; Mismatches 325; Indels 37; Gaps 9;

Qy 268 CATTCCTCTCTCTCCAAACACTTCTATTTCTTGATTTCTATCTTGCTGGCTCATTTTAAC 327
Db 28221 CATTCATTTAAATAATACATTTCTAATTTATTTTGGACTCTCTCTTGGACTATGATTA 28280
Qy 328 TCAGTAGTGAGTTGTTGGTTTCATTAAGTTTGAAGTTTCTGTTTCTGTTGTTGT 387
Db 28281 TTAGAAGTGTTGTTATTTAGTTTCCAGATATCTAGGCATTTTCCAGAG---ATCTTCTGT 28337
Qy 388 TGTGTTATCTAGATTTTAAGCTGTGGTGCAGATAGACATAGATTT-----437
Db 28338 TATTAATTTCTAATTTAAATTCATTAATAGTCATAGATATACITTTGATATATCTTACTA 28397
Qy 438 ATTTCAATGTCTTTATCTCTGAGACTTCTGTTTGGTTTGAATATGATTTCAATTTTG 497
Db 28398 ACTTGAATTTCTTTAAATTTTCTGAAACTTTGTTTATGGCCCAAGATATGCTCTACTTG 28457
Qy 498 G--AGAGTTTCATAGGGTGTGCAAGAGAGTACAGT-CTTTGTTGTTTCTGGTGAATAGT 554
Db 28458 GTAAATGCTGTATGCTTTTGGAAAGAAAGTATATTCCTCTCTTGTGGTGGAGTCT 28517

Qy 555 CTGTAATATCTCTAGGTCCACTTGGTTTATGACATCAGTTAGTCCAGCAATTTCTCTGT 614
Db 28518 CTATAAATATCAATTTAGGTCAAGTTGGTTGATAGTTATTCATATCTCTATATCTTG 28577
Qy 615 TTCGTTTTTGTGAGATGACCTAACTGTTGGAGAGATGGGTATTGCAAGTAGCCCACT 674
Db 28578 CTGATCTCTCTTCTTCTTATCTATCAATTAATGAAGATGGGCTTTGAAA-----28628
Qy 675 ATCTGCTGTGAGGTCATATGTCATTTTAGCTGTAGCTGTGCTTGTGTTTATGAACTTGG 734
Db 28629 -----TCTGCAATTTCTATTTCTACTTGAAGTCTGTGAGTTTACTTCAGGTATTTG 28683
Qy 735 GTGCAATTTGTTGGTGCATAGACATTTAAGAAATGCAATGCTCTCTTGTGGATTTTC 793
Db 28684 AAGCTCTGTATTGGGTACATAAGCGTTTAGAAATGTTATGCTCCCTTGATGAATGACC 28743
Qy 794 CTTTCATGCCATGDPAGTATTTCTCCCAATCTCATCTGCTTGTAGTTTGGGTTTAACTTA 853
Db 28744 CTTTCATCACATGAATAATACTTTTACTCTTGGAAATTAATTTTGTCTTGAATCTA 28803
Qy 854 -TTAGTCAGATATTAAAAATGACTGTATCGGCTTGCTTCTTAGGGCCATTTGCTTAGAATA 912
Db 28804 CTTTATCTCATATTAATATGACATTTTCAGATTTCTTTTGTAGTAGTGTAGCATGTATA 28863
Qy 913 --TCCTTTCCATCCTTTTACTCTAAGGTGATGCTATCCATGCTAGGTGTTGCTTTTGG 970
Db 28864 ACTTTTTTCCATACCTTTTAATCTATTTCTAACT---TTGGAGTTTAAATGTTTCTTAT 28920
Qy 971 ATGCAGCAGTAGGATGATCTTTGTTTTCATATCCATCTCTGTACCCAGTATCTTTTCTA 1030
Db 28921 AGCGAGTACATAGTTGGCTGTGTTGTTTATATCAATCTCGAAATCTCAATCCTTTAT 28980
Qy 1031 GAGAAATTAAG 1041
Db 28981 GGGTATTTAAG 28991

RESULT 35
US-09-525-160B-9/C
; Sequence 9, Application US/09525160B
; Patent No. 6569681
; GENERAL INFORMATION:
; APPLICANT: Ivanov, Evgenii
; TITLE OF INVENTION: METHODS OF IMPROVING HOMOLOGOUS RECOMBINATION
; FILE REFERENCE: 10278/016001
; CURRENT APPLICATION NUMBER: US/09/525,160B
; CURRENT FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 3033
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-525-160B-9

Query Match 5.1%; Score 117.1; DB 1; Length 3033;
Best Local Similarity 58.5%; Pred. No. 4.2e+02;
Matches 340; Conservative 0; Mismatches 224; Indels 17; Gaps 8;

Qy 379 TCTTGTGTTGTTGTTATCTAGATTTAAGCTGTGTGCTCAGATAGGACATAGACTATTA 438
Db 2079 TGTATTTATTTGGGACTTCTTTTCTTTTCTTAGTCTGTGTGAGAGAAACGTTGATATTA 2020
Qy 439 TTTCAATTTGCTTTTATCTGCGAGACTTGTCTTTTGTGAAATATGTTATTCATTTTGG 498
Db 2019 TTTCACTTTTCTTAAATTTTAAATTTTGTGACCTAGCATTTGGTATATCTCTTG 1960
Qy 499 AGAGTT--CATAGGTGCTGCAAGAGGTACAGTCTTTGTGTTTGG-TGAATAGTC 555
Db 1959 AGAATTTCCATGCTGCTGAGGAGAGAAATGTTATTTACAGCTGTTGGATCAATATTC 1900
Qy 556 TGTAAATATCT-CTAGGTCCACTTGGTTTATGACATCAGTTAGTCTCCAGCAATTTCTCTGT 614

Db 1899 TGTAAATATCTATGAGTCCATTTGACCTATGGTGCAGATGAATTTGATGTTTG----- 1845
Qy 615 TTCTGTTTTTTTGTGAGATGACCTAACTGTTGGAGAGAAATGGGTATTGAAGTAGCCCACT 674
Db 1844 TTAATTTTCTGCTGGATGATCTGTCCAAATGCTGAAGTGGGTGTTGAAGTCTCCAGTT 1785
Qy 675 ATCTGTGTGAG-----GTCAATATGTGATTTTGTAGCTGTGCTGTGCTTTTATGAA 729
Db 1784 AATATTGTATGGGGTCTGTCTATCTCTCTTTGGCTCTAAATAATTTGCTTTATATA 1725
Qy 730 CTTGGGTGACATTTGTTTGGTGCATAGACATTAAGAATTCGAATGCTCTTTGGTGGAT 789
Db 1724 CTTGAGTACTCCAGTATGGTGNATATATTTAGAAATTTGAATGCTCTTTGCTAAAT 1665
Qy 790 TTTTCTCT--TTGATGCTATGATATTTCTCCCAATCTCATCTGCTTAGTTTGGGTTAA 848
Db 1664 GGACCCCTTTATCACTATAAATGACCTTTCTTCTCTCTTTTCATAGCTTTTGTCTTGCA 1605
Qy 849 GTCTA--TTAGTCAGATTAATTAAGTATGCAATGCAATGCTCTTTAGGCACTTTGCTTA 907
Db 1604 GTCTAATTTGCTGATATAAATGAGTCTCTGCTCTCTTTTGTGTTTCCATAGGCATA 1545
Qy 908 GAATATCTTTTCCATCCCTTTACTCTAAGGTGATGCTATC 948
Db 1544 GAATAT--TTTCCATCCCTTTACCTTCAGTCTATATGTGC 1505
RESULT 36
US-09-525-160B-4/c
; Sequence 4, Application US/09525160B
; Patent No. 6569681
; GENERAL INFORMATION:
; APPLICANT: Ivanov, Evgenii
; TITLE OF INVENTION: METHODS OF IMPROVING HOMOLOGOUS RECOMBINATION
; FILE REFERENCE: 10278/016001
; CURRENT APPLICATION NUMBER: US/09/525,160B
; CURRENT FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3213
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-525-160B-4
Query Match 5.1%; Score 117.1; DB 1; Length 3213;
Best Local Similarity 58.5%; Pred. No. 4e+02;
Matches 340; Conservative 0; Mismatches 224; Indels 17; Gaps 8;
Qy 379 TGTGTTGTTGTTGTTTATCTAGATTTAAAGCTGTGGTGCAGATAGACATAGATATTA 438
Db 2079 TGATTTATTTGGACATCTTTTCTTAGTCTGTGTGAGAGAAAACGTTTGATATTA 2020
Qy 439 TTCAATGTCCTTTATCTGTGAGACTTGTCTTTGTTTGAATATGATATCAATTTGG 498
Db 2019 TTTCACTTTTCTTAAATTTAAGATTTGTTTGTGACCTAGCAATTTGGTATATCCTG 1960
Qy 499 AGAGTTT--CATAGGTGCTGACAAGAGGTACAGTCTTTGTTGTTGG--TGAAATAGTC 555
Db 1959 AGAATATCAATGCTGTAGGAGAGAAATGTGATTTCTACAGCTTTGGATCAATGTTTC 1900
Qy 556 TGTAATATCT--CPAGTCCACTTTGTTTATGACATCAGTTAGCTCCAGCAATTTCTGT 614
Db 1899 TGTAATATCTACTAGGTCCATTTGACCTATGGTGCAGATGAATTTGATGTTG----- 1845
Qy 615 TTCTGTTTTTTTGTGAGATGACCTAACTGTTGGAGAGAAATGGGTATTGAAGTAGCCCACT 674
Db 1844 TTAATTTTCTGTCTGGATGATCTGTCCAAATGCTGGAAGTGGGTGTTGAAGTCTCCAGTT 1785
Qy 675 ATCTGTGTGAG-----GTCAATATGTGATTTTGTAGCTGTAGCTTTGTTTATGAA 729
Db 1784 AATATTGTATGGGGTCTGTCTATCTCTCTTTGGCTCTAATAATTTGCTTTATATA 1725

Qy 730 CTTGGGTGACATTTGTTTGGTGCATAGACATTAAGAATTCGAATGCTCTCTTTGCTGAT 789
Db 1724 CTTGAGTACTCCAGTATTGGGTGAATATATATTTAGAATTTGTAATGCTCTCTGCTAAAT 1665
Qy 790 TTTTCTCT--TTGATGCTATGATTAATTTCTCCCAATCTCATCTGCTTAGTTTGGGTTAA 848
Db 1664 GGACCCCTTTATCACTATAAATGACCTTTCTTCTCTCTTTTTCATAGTATTGCTTGCA 1605
Qy 849 GTCTA--TTAGTCAGATTAATTAAGAATGACTGTATGCGCTTGTCTTTAGGCACTTTGCTTA 907
Db 1604 GTCTAATTTGCTGATATAAATGAGTCTCTGCTCTCTTTTGTGTTTCCATAGGCATA 1545
Qy 908 GAATATCTTTTCCATCCCTTTACTCTAAGGTGATGCTATC 948
Db 1544 GAATAT--TTTCCATCCCTTTACCTTCAGTCTATATGTGC 1505
RESULT 37
US-09-818-512-3/c
; Sequence 3, Application US/09818512
; Patent No. 6537780
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01192
; CURRENT APPLICATION NUMBER: US/09/818,512
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 116592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)--(116592)
; OTHER INFORMATION: n = A,T,C or G
US-09-818-512-3
Query Match 5.1%; Score 116; DB 1; Length 116592;
Best Local Similarity 51.8%; Pred. No. 12;
Matches 565; Conservative 0; Mismatches 465; Indels 60; Gaps 14;
Qy 212 ACTGCTCTCTCTCCCTTTCTCTAACACTTCTGGCCAGGAGTAGGGGCACCTACCGCAT 271
Db 100508 ATGTCTATACACTTCCCTTTAGAACCCGCTTTTGTGTATCCCATAGGTTTGGTATTT 100449
Qy 272 CCCTCTCTCTCCAAACACTTCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAG 331
Db 100448 GTTTCATTTTCATTTGCTCAAGAAATTTTAAATTTTACTTAAATTTCTTCATTG 100389
Qy 332 TAGTGAGTTGTTGGTTTCCATAAGTTTGTAGTTTTCGTGTTTCTGTTGTTGTTGTT 391
Db 100388 CCAATTTGTTTAAATTTCCATTAATTTGTACTGTTTCCAATGTTCC--TCTGATTATA 100332
Qy 392 GTATCTAGATTTAAGCTGTGGTGCAGATAGGACATAGAGTATTTTCAATTTGCTCT 451
Db 100331 GAATTTCTAGTTTATTTCCACGTGTGCAAGATA-----AATATAATTTCAATTTAA 100281
Qy 452 TTATCTGTGAGACTGTGTTTGTGTTTGTGAAATATGATPATTCAATTTTGGAGA--GTTTCATA 509
Db 100280 AAAATTTGTAAACCTTTGTTGTGACCTAAATGTATGATCAATCTCTGGAGAAATATCCCATG 100221
Qy 510 GGGTCTGCACAGAGAGGTACAGTCTTTGTTGTTTGG--TGAAATAGTCTGTAATATCT-C 567
Db 100220 TGTCTGTGAGAAAAATGTTATTTCTGCAACTGTTGGATGGAAATGTTCTGTAATATCTGT 100161
Qy 568 TAGGTCCACTTGGTTTATGACATCACTAGTCTCCAGCAATTTCTCTGTTTCTGTTTGT 627
Db 100160 TAGGTCCACTTGGTCCAGAGTACAGTCTTAATGCTGTACGTTTCTTGTGTTTCTTCTGCT 100101

628 GAGATGACCTAACTGTTGGAGAGAAAGGGTATTGAAGTAGCCACATCATCTGTGT-GTGA 686
Qy
Db 100100 GCATGATCTGTTCAATTTGAAA-----GTGTGGATGTTCCCTAGTATTACTGTATTGC 100048
Qy 687 GGTCAATATCTGATTTTAGCTGTAGCTGTCTGTTTATGAAGTGGGGTGCATCTGTGT 746
Db 100047 AGTCTATCTCTCTTTAGGTCTATTAAATATTGCTTTATATATTGGGTGCTCTGGTGT 99988
Qy 747 TTGGTGCATAGACATTAAGAAATGCAATGCTCTCTGGTGA-TTTTCCCTTTGATGCCCTA 805
Db 99987 TGGGTGATATATATTTACCATTTGTTATATCCCTCTCTGCTGCTGACCCCTTTGTCGTTA 99928
Qy 806 TGTAGTATCTTCCCAATCTCACTGCTGTAGTTTGGGTTTAACTCTATTAGTCAGATAT 865
Db 99927 TATATGGCCTGTCTTTTAAAC-ATCTTGATTTAAGTTTATTATTTATGTGATAA 99871
Qy 866 TAAATGACTGTATCGGCTGCTCTTTAGGGCCATTTGCTTAGAATATC-TTTTCCATCC 924
Db 99870 AACTATAGGCTCTCTGCTCTTTTGGTTTCCATTTGCAATTAATATCTTTTCCATC 99811
Qy 925 TTTTACTCTAAGGTGATGCTATCCAT---GGTAGGTGTCTTTTGGGANGCAGCAGTA 981
Db 99810 CTTTACATTTAGTCTGTGCTTATCTTTACAGATGAAGTGAGTTCTTTGTAGCCAGCATAT 99751
Qy 982 GGATGGATCTGTTTCCATATCCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAG 1041
Db 99750 GGATGGGTCTGTTTAAATCCATCTGCCATCTT-----AATGAA 99707
Qy 1042 ATCATTTAGTCAATGATGTTGAGAAATATCAATGAGCAGTGTGTTGGATCTTGTATC 1101
Db 99706 GAATTTAATCCATTTATATCAAAAGTTATTACTGACAGGTAAGGTTTACTAGCCATA 99647
Qy 1102 TTGACTGTGAAGTG 1161
Db 99646 CTGTTCCT-----TGTTTTGGTGTGCTTTGTAGATCTTTTCTTCCATTTTCC 99599
Qy 1162 TGCTG 1221
Db 99598 TCTCTTACTGCTTCTCTATGTAAGTATTTCTCTAGCAGGTGCTTTGATTTGTG 99539
Qy 1222 ATTTTGGCCCTGGAATTTATTTATTTATTTATTTTCTTTGAAATGGGTAACATCTTTAGA 1281
Db 99538 CTATTTATTTTGTGTATCTATATAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 99479
Qy 1282 TTGAAGTTT 1291
Db 99478 AAAAAATCTT 99469

RESULT 38

US-09-734-673-3
; Sequence 3, Application US/09734673
; Patent No. 6410294
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001020
; CURRENT APPLICATION NUMBER: US/09/734, 673
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 38564
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(38564)
; OTHER INFORMATION: n = A,T,C or G
US-09-734-673-3

Query Match 5.1%; Score 115.3; DB 1; Length 38564;
Best Local Similarity 59.1%; Pred. No. 36;
Matches 298; Conservative 0; Mismatches 197; Indels 9; Gaps 6;
Qy 298 TCTTCATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGGTTTCCATAAGT 357
Db 22517 TCTTTATTTCTATTGTTAAACCAAAATTTATACAGGAGCAGATTTTAAATTTCCATGTAT 22576
Qy 358 TTGTAAGTTTCTGTGTGTTTCTGTGTTGTTGTTGTTATCTAGATTTAAGCTGTGGTGT 417
Db 22577 TTGTATAGTTTGGAAATCTCTTTTGGAGTTGGTT---TCTAGTTTTCCTACTGTGT 22633
Qy 418 CAGATAGACATAGAGTATTTTCAATGCTTTTATCTCTCGAGACTTCTGTTTGT 477
Db 22634 CTGAGAAAGATGCTGATATGATTTTGAATTTTAAAAAT-TATTGATCTGTTTGTAGC 22692
Qy 478 GAAATATGTTTCAATTTTGA--GAGTTTCTATAGGGTCTGCAAGAGGTACAGTCT- 534
Db 22693 CAATCGTATGCCCTATGTTGGAATATATTCATGCTGATGAGAGAAAGTATATCTA 22752
Qy 535 TTGCTTTTGGTGAATAGTCTGTAATATCTCTAGTCCACTTGGTTTATGACATCAGT 594
Db 22753 TAGTTTGGGTAGGATGTTCTGTAATATCTGTAGGCTCTACTTGTGTTTGTAGATATTGT 22812
Qy 595 TAGCTCCAGCATTTCTCTGTTTCTGTTTGTGAGATGACCTAACTGTTGGAGAGAAATG 654
Db 22813 TTAAGCTATTGTTCTTTG-TTGACTTCTGCTCGATGATCTGCTAGTCTGTCTAGTG 22871
Qy 655 GGGTATTGAAGTAGCCCACTATCTGTGTG-TGAGGTCAATATGATATTTAGCTGTAGCT 713
Db 22872 GGGTATTGAAGTAGCCCACTATTTATGGGTTGCTCAATATCTCTTCCCTAGGCTTAGCA 22931
Qy 714 GTGCTGTTTATGAACCTGGGTGACATTTGTTGGTGCATAGACATTAAGAATTGCA 773
Db 22932 GTAAATGTTTATGAATTTTGAGCTCCATAGTTAGGTACACATATTTTGGATTATA 22991
Qy 774 TGTCCTCTGGTGGATTTTCTTT 797
Db 22992 TATTCTCTGTCAGATTAATCCTT 23015

RESULT 39

US-09-010-398-14
; Sequence 14, Application US/09010398
; Patent No. 5972615
; GENERAL INFORMATION:
; APPLICANT: An. Gang
; TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS, DISEASE
; TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,398
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Corder, Timothy S.
; REGISTRATION NUMBER: 38,414
; REFERENCE/DOCKET NUMBER: UROC:015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-418-3000


```
; TELEFAX: 512-474-7577
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8224 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-010-398-14

Query Match          4.8%; Score 108.3; DB 1; Length 8224;
Best Local Similarity 59.2%; Pred. No. 1.7e+02;
Matches 218; Conservative 0; Mismatches 147; Indels 3; Gaps 2;

QY 1345 TTTTATCTAGATGCTCTTTCTTCTCAACTATTGTGACAGAAAGTTTCTTAAGTGCA 1404
Db 640 TATTCTCTCTGAACGATTTATTCTCTCTCAATCTCTGAAGGATAGCTTTGCTGAGGGTA 699
QY 1405 GTAGTCTGGCGTGACATCTGTAGTCTCTTGAGTCTGTAGCAGATCTGTGAGGGCCTTC 1464
Db 700 ATGTTCTTAGCTGACAGTTTCTCTTTTAGTAATTGAATATATTA--TGCCATTCTTTC 757
QY 1465 TTACATTTTGAGTTTCTATTGGAAGTCAGGTGTAATCTAATACATCTGCCCTTATAT 1524
Db 758 CTGGCCTGTAGGTTTCTCCGAGAAATTCCTCTTAGTCAGATGGGGGTACCTTATAT 817
QY 1525 GTTAATTGGCTTTTCCCTTGCACTCTTTTAATATCTTTCTTTGTCTATATCTTTTAG 1584
Db 818 GTGACTTGA-CATTTTCTCTTCTGCTTTTAAATAATCTTTCTTTGCTTTGACITTTAG 876
QY 1585 TGATTGATTAATGACATGTGGGGAGTTTCTTTCCGTCGAATCTATTGTGGTGT 1644
Db 877 CAATGCAATTAATGCTTGGAGAGAACCCGTTTAGGTGAATTAATTAGGGTCT 936
QY 1645 GTATGCTTCTGTACCTGTAGGATCTCTTCTCAAGTTAGGAATTTCTTTT 1704
Db 937 TATAGTCTCTGACCTGGATTTCTCTCTCTCCAGATGCAAGAAGTTTCTGCTAT 996
QY 1705 GGTITCT 1712
Db 997 TATTTAT 1004

RESULT 40
US-09-366-260-14
; Sequence 14, Application US/09366260
; Patent No. 6171796
; GENERAL INFORMATION:
; APPLICANT: An, Gang
; Veltri, Robert W.
; TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,
; PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/366,260
; FILING DATE: 03-Aug-1993
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/010,398
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Corder, Timothy S.
```

```
; REGISTRATION NUMBER: 38,414
; REFERENCE/DOCKET NUMBER: UROC:015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-418-3000
; TELEFAX: 512-474-7577
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8224 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-366-260-14

Query Match          4.8%; Score 108.3; DB 1; Length 8224;
Best Local Similarity 59.2%; Pred. No. 1.7e+02;
Matches 218; Conservative 0; Mismatches 147; Indels 3; Gaps 2;

QY 1345 TTTTATCTTAGAATGCTTTTCTTCTCAACTATTGTGACAGAAAGTTTCTTAAGTGCA 1404
Db 640 TATTCTCTCTGAACGATTTATTCTCTCTCAATCTCTGAAGGATAGCTTTGCTGAGGGTA 699
QY 1405 GTAGTCTGGCGTGACATCTGTAGTCTCTTGAGTCTGTAGCAGATCTGTGAGGGCCTTC 1464
Db 700 ATGTTCTTAGCTGACAGTTTCTCTTTTAGTAATTGAATATATTA--TGCCATTCTTTC 757
QY 1465 TTACATTTTGAGTTTCTATTGGAAGTCAGGTGTAATCTAATACATCTGCCCTTATAT 1524
Db 758 CTGGCCTGTAGGTTTCTCCGAGAAATTCCTCTTAGTCAGATGGGGGTACCTTATAT 817
QY 1525 GTTAATTGGCTTTTCCCTTGCACTCTTTTAATATCTTTCTTTGTCTATATCTTTTAG 1584
Db 818 GTGACTTGA-CATTTTCTCTTCTGCTTTTAAATAATCTTTCTTTGCTTTGACITTTAG 876
QY 1585 TGATTGATTAATGACATGTGGGGAGTTTCTTTCCGTCGAATCTATTGTGGTGT 1644
Db 877 CAATGCAATTAATGCTTGGAGAGAACCCGTTTAGGTGAATTAATTAGGGTCT 936
QY 1645 GTATGCTTCTGTACCTGTAGGATCTCTTCTCAAGTTAGGAATTTCTTTT 1704
Db 937 TATAGTCTCTGACCTGGATTTCTCTCTCTCCAGATGCAAGAAGTTTCTGCTAT 996
QY 1705 GGTITCT 1712
Db 997 TATTTAT 1004

RESULT 41
US-09-702-705-604
; Sequence 604, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; NUMBER OF SEQUENCES: 30
; FILING DATE: 2000-10-30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 604
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
```

```
/ NAME/KEY: misc feature
/ LOCATION: (1)...(468)
/ OTHER INFORMATION: n = A,T,C or G
US-09-702-705-604

Query Match 4.7%; Score 106.2; DB 1; Length 468;
Best Local Similarity 57.7%; Pred. No. 2.9e+03;
Matches 239; Conservative 0; Mismatches 171; Indels 4; Gaps 3;

381 TTGTTGTTGTTTATCTAGATTAAAGCTGTGGTGCAGATAGACATAGAGTATTATT 440
17 TTCTTAATCCCTGAGTTCTGGNTTGAATTCACACTGTGCTGAGAGATAGTTTGTATAATT 76
441 TCAATTGCTTTTATCTGTCGAGACTTCTGTTTGAATAATATCTATCAATTTTGGAG 500
77 TCTGTTCTTTTACACTTACTGAGGAGAGCTTTACTTCCAGTAAGTGTGCTCGAATTTGGAA 136
501 A--GTTTCATAGGGTCTGCAAGAAGGTACAGTCT--TTGTGTTTGGTGAATAGTCTG 557
137 TAGGTGTGGTGTGCTGCTGAAAGAAGTATATATCTGTTGATTTGGGGTGGAGAGTTCTG 196
558 TAAATATCTCTAGGTCACACTTGGTTTATGACATCAGTTAGTCCAGCAATTTCTCTGTTTC 617
197 TANATGTCTATTAGTCCGCTTGGTGCAGAGTTGAGTTCAATTCCTGGATAGCCTTGTGA 256
618 GTTTTTTGTGAGATGACCTAACTGTTGGAGAGAAATGGGTATTGAAGTAGCCCACTATC 677
257 ACTTTCTGCTCGTTGATCTGCTAAATGTTGACAGTGGGTGGAAGTCTCCCATTAAT 316
678 TGTGTGTGAG-GTCAATATGTCATTTTATGATCAGTTAGTCTGCTGTTTATGAAGTGGGT 736
317 ATTGTGTGGGAGCTTAAAGTCTCTTTGTAGTCACTAAGGAGCTTGCCTTTATGAATCTGGGT 376
737 GACATTTGTTTGGTGCATAGACATTAAGAATTGCAATGCTCTCTTGGTGGATT 790
377 GTCCTGCAATGGGTGGACATATATTTAGGACAGCAGCTCTTCTTGTGAATT 430

RESULT 43
US-09-614-124B-604
/ Sequence 604, Application US/09614124B
/ Patent No. 6630574
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C9
/ CURRENT APPLICATION NUMBER: US/09/614,124B
/ CURRENT FILING DATE: 2001-07-11
/ NUMBER OF SEQ ID NOS: 1668
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 604
/ LENGTH: 468
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)...(468)
/ OTHER INFORMATION: n = A,T,C or G
US-09-614-124B-604

Query Match 4.7%; Score 106.2; DB 1; Length 468;
Best Local Similarity 57.7%; Pred. No. 2.9e+03;
Matches 239; Conservative 0; Mismatches 171; Indels 4; Gaps 3;

381 TTGTTGTTGTTTATCTAGATTAAAGCTGTGGTGCAGATAGACATAGAGTATTATT 440
17 TTCTTAATCCCTGAGTTCTGGNTTGAATTCACACTGTGCTGAGAGATAGTTTGTATAATT 76
441 TCAATTGCTTTTATCTGTCGAGACTTCTGTTTGAATAATATGTAATCAATTTTGGAG 500
77 TCTGTTCTTTTACACTTACTGAGGAGAGCTTTACTTCCAGTAAGTGTGCTCGAATTTGGAA 136
501 A--GTTTCATAGGGTCTGCAAGAAGGTACAGTCT--TTGTGTTTGGTGAATAGTCTG 557
137 TAGGTGTGGTGTGCTGCTGAAAGAAGTATATATCTGTTGATTTGGGGTGGAGAGTTCTG 196
558 TAAATATCTCTAGGTCACACTTGGTTTATGACATCAGTTAGTCCAGCAATTTCTCTGTTTC 617
197 TANATGTCTATTAGTCCGCTTGGTGCAGAGTTGAGTTCAATTCCTGGATAGCCTTGTGA 256
618 GTTTTTTGTGAGATGACCTAACTGTTGGAGAGAAATGGGTATTGAAGTAGCCCACTATC 677
257 ACTTTCTGCTCGTTGATCTGCTAAATGTTGACAGTGGGTGGAAGTCTCCCATTAAT 316
678 TGTGTGTGAG-GTCAATATGTCATTTTATGATCAGTTAGTCTGCTGTTTATGAAGTGGGT 736
317 ATTGTGTGGGAGCTTAAAGTCTCTTTGTAGTCACTAAGGAGCTTGCCTTTATGAATCTGGGT 376
737 GACATTTGTTTGGTGCATAGACATTAAGAATTGCAATGCTCTCTTGGTGGATT 790
377 GTCCTGCAATGGGTGGACATATATTTAGGACAGCAGCTCTTCTTGTGAATT 430

RESULT 42
US-09-736-457-604
/ Sequence 604, Application US/09736457
/ Patent No. 6509448
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Lodes, Michael A.
/ APPLICANT: Fanger, Gary
/ APPLICANT: Vedvick, Tom
/ APPLICANT: Carter, Darrick
/ APPLICANT: Retter, Marc
/ APPLICANT: Mannion, Jane
/ APPLICANT: Fan, Liqun
/ APPLICANT: Wang, Aijun
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
/ TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.478C15
/ CURRENT APPLICATION NUMBER: US/09/736,457
/ CURRENT FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 1864
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 604
/ LENGTH: 468
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)...(468)
/ OTHER INFORMATION: n = A,T,C or G
US-09-736-457-604

Query Match 4.7%; Score 106.2; DB 1; Length 468;
Best Local Similarity 57.7%; Pred. No. 2.9e+03;
Matches 239; Conservative 0; Mismatches 171; Indels 4; Gaps 3;
```


us10664775-4.rni

Wed Aug 25 15:11:59 2004

Search completed: August 25, 2004, 08:12:44
Job time : 3688 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	2279	100.0	2279	1	US-10-664-775-4	Sequence 4, Appli
2	2279	100.0	2715	1	US-10-664-775-1	Sequence 1, Appli
3	2251	98.8	2267	1	US-10-664-775-5	Sequence 5, Appli
4	2003	87.9	2003	1	US-10-664-775-3	Sequence 3, Appli
5	626.9	27.5	4129	1	US-10-087-192-577	Sequence 577, App
	610.3	26.8	193853	1	US-10-087-192-1663	Sequence 1663, App
C 7	591.499	26.0	19619	1	US-10-369-224-129	Sequence 129, Appl
C 8	590.199	25.9	178825	1	US-10-087-192-1945	Sequence 1945, Ap
C 9	588.5	25.8	6335	1	US-10-368-934-580	Sequence 580, Appl
C 10	588.5	25.8	6335	1	US-10-388-934-685	Sequence 685, Appl
C 11	588.5	25.8	6335	1	US-10-388-934-790	Sequence 790, Appl
C 12	588.5	25.8	6335	1	US-10-191-803-720	Sequence 720, Appl
C 13	588.3	25.8	29604	1	US-10-374-077-207	Sequence 207, Appl
C 14	587.8	25.8	289150	1	US-10-322-281-115	Sequence 115, Appl
C 15	585.3399	25.7	225883	1	US-10-175-523-57	Sequence 57, Appl
C 16	585.1	25.7	173808	1	US-10-003-806-10	Sequence 10, Appl
C 17	580.301	25.5	120239	1	US-10-322-281-271	Sequence 271, Appl
C 18	578.299	25.4	173808	1	US-10-003-806-10	Sequence 10, Appl
C 19	578.2	25.4	8048	1	US-10-388-934-223	Sequence 223, Appl
C 20	578.2	25.4	8048	1	US-10-388-934-224	Sequence 224, Appl
C 21	578.2	25.4	8048	1	US-10-388-934-682	Sequence 682, Appl
C 22	578.2	25.4	8048	1	US-10-388-934-720	Sequence 720, Appl
C 23	578.2	25.4	8048	1	US-10-388-934-785	Sequence 785, Appl
C 24	578.2	25.4	8048	1	US-10-388-934-808	Sequence 808, Appl
C 25	574.999	25.2	116858	1	US-10-087-192-547	Sequence 547, Appl
C 26	573.4	25.2	200400	1	US-10-087-192-1033	Sequence 1033, Ap
C 27	572.9	25.1	225883	1	US-10-175-523-57	Sequence 57, Appl
C 28	570.9	25.1	95982	1	US-10-087-192-2029	Sequence 2029, Appl
C 29	567.5	24.9	251364	1	US-10-175-523-58	Sequence 58, Appl
C 30	567.5	24.9	251364	1	US-10-175-523-61	Sequence 61, Appl
C 31	567.5	24.9	251364	1	US-10-175-523-79	Sequence 79, Appl
C 32	566.4	24.9	175338	1	US-10-087-192-1861	Sequence 1861, Appl
C 33	565.399	24.8	51259	1	US-10-374-077-209	Sequence 209, Appl


```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2715
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-10-664-775-1

Query Match      100.0%; Score 2279; DB 1; Length 2715;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 2279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GATCACCCTCTAGTGAAGGTGGGGTCTCAGGCTCCAAATGGTGTGTGATGGTAGAG 60
DB      437  GATCACCCTCTAGTGAAGGTGGGGTCTCAGGCTCCAAATGGTGTGTGATGGTAGAG 496
QY      61  TATCTCATACAGAGATAGCACTAGATGCTCTCTGGGACATAGATGAAGCTTTCAGAGAG 120
DB      497  TATCTCATACAGAGATAGCACTAGATGCTCTCTGGGACATAGATGAAGCTTTCAGAGAG 556
QY      121  ACTTCATAATATATTTCTTCAAGCCCTGCTGGCAATACCTTCGGGGCTGCTGCCCTTC 180
DB      557  ACTTCATAATATATTTCTTCAAGCCCTGCTGGCAATACCTTCGGGGCTGCTGCCCTTC 616
QY      181  TCCCTGCTCTGATTCCTAGGGTGAGGGTTACCACTGCTCTCTCTCTCTCTCTCTCTCT 240
DB      617  TCCCTGCTCTGATTCCTAGGGTGAGGGTTACCACTGCTCTCTCTCTCTCTCTCTCTCT 676
QY      241  TTCTGGGCCAGGTAGGGGCACTACCGCAATTCCTCTCTCTCTCTCTCTCTCTCTCTCT 300
DB      677  TTCTGGGCCAGGTAGGGGCACTACCGCAATTCCTCTCTCTCTCTCTCTCTCTCTCTCT 736
QY      301  TGATTTCTATCTTGCTCATTTTAACTCAGTAGTAGTGTGTTGGTTTCCATAAGTTTG 360
DB      737  TGAATTTCTATCTTGGCTCATTTTAACTCAGTAGTAGTGTGTTGGTTTCCATAAGTTTG 796
QY      361  TAAATTTCTGTGTTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 420
DB      797  TAAATTTCTGTGTTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 856
QY      421  ATAGACATAGATATTTCAATGCTTTTATCTGTCGAGACTTGTCTGTTGTTGTTGTTG 480
DB      857  ATAGACATAGATATTTCAATGCTTTTATCTGTCGAGACTTGTCTGTTGTTGTTGTTG 916
QY      481  ATATGTATTCATTTTGGAGAGTTTTCATAGGGTGTGTCAGAAAGGTACAGTCTTTGTT 540
DB      917  ATATGTATTCATTTTGGAGAGTTTTCATAGGGTGTGTCAGAAAGGTACAGTCTTTGTT 976
QY      541  TTTGGTGAATAGTCTGTAATATCTAGGTCCACTTGGTTTATGACATCAGTAGTCTC 600
DB      977  TTTGGTGAATAGTCTGTAATATCTAGGTCCACTTGGTTTATGACATCAGTAGTCTC 1036
QY      601  CAGCAATTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 660
DB      1037  CAGCAATTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT 1096
QY      661  TGAAGTAGCCCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
DB      1097  TGAAGTAGCCCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1156
QY      721  TTTTATGACCTTGGGTGACATTTGTTTGGTGGCAATAGACATTAAGAAATGCAATGCTC 780
DB      1157  TTTTATGACCTTGGGTGACATTTGTTTGGTGGCAATAGACATTAAGAAATGCAATGCTC 1216
QY      781  TTGTGGATTTTCTTTGATGCTATGTAGTATTTCTCCCAATCTCATCTGCTTAGTTT 840
DB      1217  TTGTGGATTTTCTTTGATGCTATGTAGTATTTCTCCCAATCTCATCTGCTTAGTTT 1276
QY      841  GGGTTTATAGTCTATTAGTCAGATATTAATAAGACGTGTATCGCTTCTCTTAGGGCAT 900
DB      1277  GGGTTTATAGTCTATTAGTCAGATATTAATAAGACGTGTATCGCTTCTCTTAGGGCAT 1336
QY      901  TTGCTTAGAATATCTTTTCCATCTTTTACCTTAAGGTGATGTCTATCCATGGTAGGTTG 960
DB      1337  TTGCTTAGAATATCTTTTCCATCTTTTACCTTAAGGTGATGTCTATCCATGGTAGGTTG 1396
QY      961  TCTTTTGTGATGACAGCAGTAGATGGATCTTTGTTTTCATATCCATCTCTGTACCCAGTA 1020
DB      1397  TCTTTTGTGATGACAGCAGTAGATGGATCTTTGTTTTCATATCCATCTCTGTACCCAGTA 1456
QY      1021  TCTTTTCTAGAGAAATTAAGATCATTTGATGATCTTGAAGATTTAGAGATTTCAATGAGCAG 1080
DB      1457  TCTTTTCTAGAGAAATTAAGATCATTTGATGATCTTGAAGATTTAGAGATTTCAATGAGCAG 1516
QY      1081  TGTTTTGTGATTTCTTTGATTTCTTTGCACTTGTGAAGTGTGTGTGTGTGTGTGTGTGTGT 1140
DB      1517  TGTTTTGTGATTTCTTTGATTTCTTTGCACTTGTGAAGTGTGTGTGTGTGTGTGTGTGTGT 1576
QY      1141  GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
DB      1577  GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1636
QY      1201  TGTGTTCTCTCCCTCTTTTGTGATTTTGGCTTGGAAATTAATTAATTAATTAATTAATTAAT 1260
DB      1637  TGTGTTCTCTCCCTCTTTTGTGATTTTGGCTTGGAAATTAATTAATTAATTAATTAATTAAT 1696
QY      1261  AATGTGGGTAAACATCTTTAGATTAAGTATTTCTCTAGCCTTCTTTAGGCTGTGATTTG 1320
DB      1697  AATGTGGGTAAACATCTTTAGATTAAGTATTTCTCTAGCCTTCTTTAGGCTGTGATTTG 1756
QY      1321  AAGATAGATATTTCTTTTACATCTGATTTTATCTTAGAATGCTTTCTTCTTCCAACTATTG 1380
DB      1757  AAGATAGATATTTCTTTTACATCTGATTTTATCTTAGAATGCTTTCTTCTTCCAACTATTG 1816
QY      1381  TGA CAGAAAGTTTTTCTAAAGTCAGTAGTCTGGCCTGACATCTGTAGTCTCTTGGAGTCT 1440
DB      1817  TGA CAGAAAGTTTTTCTAAAGTCAGTAGTCTGGCCTGACATCTGTAGTCTCTTGGAGTCT 1876
QY      1441  GTAGCACATCTGTGAGGGGCTTCTTACATTTTGTAGTCTTCTTCCCTTGGCATCTTTTAATAT 1500
DB      1877  GTAGCACATCTGTGAGGGGCTTCTTACATTTTGTAGTCTTCTTCCCTTGGCATCTTTTAATAT 1936
QY      1501  ATTCTAATACATCTGCTCTTATATGTTAATTTGGTCTTTTCCCTTGGCATCTTTTAATAT 1560
DB      1937  ATTCTAATACATCTGCTCTTATATGTTAATTTGGTCTTTTCCCTTGGCATCTTTTAATAT 1996
QY      1561  TCTTTCTTTGTTCTATACCTTTTAGTGATTTGATTAATATGACATCTGTTGGGAGTTCTTTT 1620
DB      1997  TCTTTCTTTGTTCTATACCTTTTAGTGATTTGATTAATATGACATCTGTTGGGAGTTCTTTT 2056
QY      1621  CCGGTCCAACTATTTTGGTGTGTTTGTATGCTTCTTGTACCTTGTAGAGGATCTCTTTTCTC 1680
DB      2057  CCGGTCCAACTATTTTGGTGTGTTTGTATGCTTCTTGTACCTTGTAGAGGATCTCTTTTCTC 2116
QY      1681  AAGGTAGGAAATTTTCTTTTGGTGTGTTTGTGAAATATTTTCCCTGCTTTTGACCTG 1740
DB      2117  AAGGTAGGAAATTTTCTTTTGGTGTGTTTGTGAAATATTTTCCCTGCTTTTGACCTG 2176
QY      1741  CTTCTCTCCCTTCCCTTATCTCTTTTGGTGTGATGATGCTCTGCTTCCCTGATGT 1800
DB      2177  CTTCTCTCCCTTCCCTTATCTCTTTTGGTGTGATGATGCTCTGCTTCCCTGATGT 2236
QY      1801  TTTATGCTCGATATTTTGTAGACTTAACATTTTCTTTGACCAAGGTATCCATTTCTTCTA 1860
DB      2237  TTTATGCTCGATATTTTGTAGACTTAACATTTTCTTTGACCAAGGTATCCATTTCTTCTA 2296
QY      1861  TCTTGCTTCACTGCTGAGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1920
DB      2297  TCTTGCTTCACTGCTGAGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2356
QY      1921  TCTCTGAGGTTCTGTTGGGTTCTTAATTTTTCATTTTCCAGATTTTCCCTCAGTTGGGT 1980
DB      2357  TCTCTGAGGTTCTGTTGGGTTCTTAATTTTTCATTTTCCAGATTTTCCCTCAGTTGGGT 2416
QY      1981  TTTGTTTATTAATTTCTATTTCACCTTTCAGGCTCCTGAAATGTTTACTCATTTTCTCCTC 2040
DB      2417  TTTGTTTATTAATTTCTATTTCACCTTTCAGGCTCCTGAAATGTTTACTCATTTTCTCCTC 2476
```

QY 2041 AGTATTTTACATTTTCATAGGTTTCTTTAATGAATTAATCAATTCCTCTTCAAGGACCTT 2100
DB 2477 AGTA-TTACATTTTCATAGGTTTCTTTAATGAATTAATCAATTCCTCTTCAAGGACCTT 2536
QY 2101 TTATGAATTCATAAATATGATGTTAAGTCTCTTGGCTGTCCTTCAGGTATGTTGCAATTC 2160
DB 2537 TTATGAATTCATAAATATGATGTTAAGTCTCTTGGCTGTCCTTCAGGTATGTTGCAATTC 2596
QY 2161 TCAGGGCCCTATTGTAATAGGTTTATAGCAGGACATATGTCCTGGTGTGTTAATGTCGT 2220
DB 2597 TCAGGGCCCTATTGTAATAGGTTTATAGCAGGACATATGTCCTGGTGTGTTAATGTCGT 2656
QY 2221 GTTTTGTCTTTGGCATATAGACGGCTGAGTTTGGGATGATGTAATTCATAGGTCGTGAT 2279
DB 2657 GTTTTGTCTTTGGCATATAGACGGCTGAGTTTGGGATGATGTAATTCATAGGTCGTGAT 2715

RESULT 3
US-10-664-775-5
; Sequence 5, Application US/10664775
; Publication No. US2004011576A1
; GENERAL INFORMATION:
; APPLICANT: Simesen, Ruth B
; APPLICANT: Pedersen, Anette A
; APPLICANT: Faisst, Steffen
; APPLICANT: Jensen, Jan J
; APPLICANT: Wellguy, Dietmar
; TITLE OF INVENTION: Method for Making Recombinant Proteins
; FILE REFERENCE: 6448.200-US
; CURRENT APPLICATION NUMBER: US/10/664,775
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: Danish Application PA 2002 01384
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 60/416,566
; PRIOR FILING DATE: 2002-10-07
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 5
; LENGTH: 2267
; TYPE: DNA
; ORGANISM: Cricetus griseus
US-10-664-775-5

Query Match 98.8%; Score 2251; DB 1; Length 2267;
Best Local Similarity 99.5%; Pred. No. 1.4e-06;
Matches 2267; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

QY 1 GATCACTCCTCTAGTGAAGGTGGGGTCTGAGGCTCCAAATGGTGTGATGTTGATGTTAGAG 60
DB 1 GATCACTCCTCTAGTGAAGGTGGGGTCTGAGGCTCCAAATGGTGTGATGTTGATGTTAGAG 60
QY 61 TATCTCATACAGAGATAGCACTAGATGCTCTCTGGGACATAGTAAGCTTCCAGAGAG 120
DB 61 TATCTCATACAGAGATAGCACTAGATGCTCTCTGGGACATAGTAAGCTTCCAGAGAG 120
QY 121 ACTTCATAATATATTTCTTGAAGCCTCTGCTGGCAATATCTTCTGGGGCTGCTGCCCTTTC 180
DB 121 ACTTCATAATATATTTCTTGAAGCCTCTGCTGGCAATATCTTCTGGGGCTGCTGCCCTTTC 180
QY 181 TCCCTGTCTGATTCCTAGGTTAGGGTTACCACTGCTCTCTCTCCCTTCTCTTAACAC 240
DB 181 TCCCTGTCTGATTCCTAGGTTAGGGTTACCACTGCTCTCTCTCCCTTCTCTTAACAC 240
QY 241 TTCTGGGCGAGGTAGGGGCACTACCGCATTCCTCTCTCTTCCAAACACTTCTTAATTTCT 300
DB 241 TTCTGGGCGAGGTAGGGGCACTACCGCATTCCTCTCTCTTCCAAACACTTCTTAATTTCT 300
QY 301 TGATTTCTATCTGGCTCAATTTTAACTCAGTAGTAGTGTGTTTCCATAAGTTTG 360
DB 301 TGATTTCTATCTGGCTCAATTTTAACTCAGTAGTAGTGTGTTTCCATAAGTTTG 360
QY 361 TAAGTTTCTGTGTTTCTGTGTTGTTGTTATCTAGATTTAAGCTGTGGGTGTCAG 420

DB 361 TAAGTTTCTGTGTTTCTGTGTTGTTGTTATCTAGATTTAAGCTGTGGGTGTCAG 420
QY 421 ATAGGACATAGAGTATTAATTTCAATTTGCTTTATCTGCGAGACTTGTCTTTGTTTCAA 480
DB 421 ATAGGACATAGAGTATTAATTTCAATTTGCTTTATCTGCGAGACTTGTCTTTGTTTCAA 480
QY 481 ATATGTAATCAATTTTGGAGAGTTTCAATAGGTTGCTGACAAGAGGTACAGTCTTTGTGT 540
DB 481 ATATGTAATCAATTTTGGAGAGTTTCAATAGGTTGCTGACAAGAGGTACAGTCTTTGTGT 540
QY 541 TTTTGTGAATAGTCTGTAATATCTCTAGGTCACACTTGGTTTATAGACATCAGTTAGCTC 600
DB 541 TTTTGTGAATAGTCTGTAATATCTCTAGGTCACACTTGGTTTATAGACATCAGTTAGCTC 600
QY 601 CAGCAATTTCTGTGTTTCTGTGTTTCTGAGATGACCTAACTGTTGGAGAGAAATGGGTAT 660
DB 601 CAGCAATTTCTGTGTTTCTGTGTTTCTGAGATGACCTAACTGTTGGAGAGAAATGGGTAT 660
QY 661 TGAAGTAGCCCACTATCTGTGTTGAGGTCAATATGATTTAGCTGTAGCTGTGCTTTG 720
DB 661 TGAAGTAGCCCACTATCTGTGTTGAGGTCAATATGATTTAGCTGTAGCTGTGCTTTG 720
QY 721 TTTTATGAATTTGGGTGACATTTGTTGGTGCATAGACATTAAGAAATGCAATGCTCCTC 780
DB 721 TTTTATGAATTTGGGTGACATTTGTTGGTGCATAGACATTAAGAAATGCAATGCTCCTC 780
QY 781 TTGGTGAATTTTCCCTTTGATGCTATGATTTTCCCAATCTCATCTGCTTGTAGTTT 840
DB 781 TTGGTGAATTTTCCCTTTGATGCTATGATTTTCCCAATCTCATCTGCTTGTAGTTT 840
QY 841 GGGTTTAACTTATAGTACAGATATTAAGTACATGATCGGCTTCTCTTAGGCGCAT 900
DB 841 GGGTTTAACTTATAGTACAGATATTAAGTACATGATCGGCTTCTCTTAGGCGCAT 900
QY 901 TTGCTTAGAATATCTTTTCCATCTTTTACCTAAGGTGATGCTATCCATGTTAGTTG 960
DB 901 TTGCTTAGAATATCTTTTCCATCTTTTACCTAAGGTGATGCTATCCATGTTAGTTG 960
QY 961 TCTTTTGGATGCGAGCAGTAGGATGATCTGTTTTCATATCCATCTCTGTATCCAGTA 1020
DB 961 TCTTTTGGATGCGAGCAGTAGGATGATCTGTTTTCATATCCATCTCTGTATCCAGTA 1020
QY 1021 TCTTTTCTAGAGAAATTAAGATCAATGATGATGATGTTGAGAAATTAATCAATGAGCAG 1080
DB 1021 TCTTTTCTAGAGAAATTAAGATCAATGATGATGATGTTGAGAAATTAATCAATGAGCAG 1080
QY 1081 TGTGTTGAGATCTTGTGTTATCTTGGACATTTGGAAGTGTGTTGTTGTTGTTGTTGTT 1140
DB 1081 TGTGTTGAGATCTTGTGTTATCTTGGACATTTGGAAGTGTGTTGTTGTTGTTGTTGTT 1128
QY 1141 GTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
DB 1129 GTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1188
QY 1201 TGTGTTCTCTCCCTCTTTTGAATTTTGGCTTGGAAATTAATTAATTAATTAATTAATTT 1260
DB 1189 TGTGTTCTCTCCCTCTTTTGAATTTTGGCTTGGAAATTAATTAATTAATTAATTTCTTG 1248
QY 1261 AATGTTGGGTAAATCTTATAGATTTGAGTTTCTCTAGCTTCTTATAGGTTCTGATTTG 1320
DB 1249 AATGTTGGGTAAATCTTATAGATTTGAGTTTCTCTAGCTTCTTATAGGTTCTGATTTG 1308
QY 1321 AAGATAGATTAATCTTACATCTGATTTTATCTTAGAATGCTCTTCTTCTTCTTCTTCT 1380
DB 1309 AAGATAGATTAATCTTACATCTGATTTTATCTTAGAATGCTCTTCTTCTTCTTCTTCT 1368
QY 1381 TCACAGAAAGTTTCTTAAAGTACAGTGTGGCTGACATCTGTAGTCTCTTGGAGTCT 1440
DB 1369 TCACAGAAAGTTTCTTAAAGTACAGTGTGGCTGACATCTGTAGTCTCTTGGAGTCT 1428
QY 1441 GTAGACATCTGTGAGGCGCTTCTTACATTTTGAATTTTCTTATTTGAAAGTCTGAGTGA 1500

Db 1429 GTAGCACATCTGTGCAGGCGCTCTTACATTTTGAGTTTCTATTGGAAAAGTCAGGTGTA 1488
Qy 1501 ATTCTAATACATCTGCGCTTTATATAGTTAAATGGTCTTTTTCCTTCCTGCACTTTTAATAT 1560
Db 1489 ATTCTAATACATCTGCGCTTTATATAGTTAAATGGTCTTTTTCCTTCCTGCACTTTTAATAT 1548
Qy 1561 TCTTTCTTTGTTCTATACATCTTTAGTATTTGATTTATGATTTATGACATGTCGGAGTTTCTTTT 1620
Db 1549 TCTTTCTTTGTTCTATACATCTTTAGTATTTGATTTATGATTTATGACATGTCGGAGTTTCTTTT 1608
Qy 1621 CCGGTCCAAATCTATTGTGTTGTTTGTATGCTTCTTTGACCTTTGATAGGCACTCTTTCTC 1680
Db 1609 CCGGTCCAAATCTATTGTGTTGTTTGTATGCTTCTTTGACCTTTGATAGGCACTCTTTCTC 1668
Qy 1681 AAGGTAGAAATTTTCTTTTGTGTTTGTGTTTCTTGAATAATTTTCCCTGCTTTTGACCTG 1740
Db 1669 AAGGTAGAAATTTTCTTTTGTGTTTGTGTTTCTTGAATAATTTTCCCTGCTTTTGACCTG 1728
Qy 1741 CTTCTTCCCTCTCTCTATTTCCCTTTGTTTGTGATGCTCTGCTGCTTCTGCTGATGT 1800
Db 1729 CTTCTTCCCTCTCTCTATTTCCCTTTGTTTGTGATGCTCTGCTGCTTCTGCTGATGT 1788
Qy 1801 TTATGCTGATTAATTTTAGACTTAAATTTTCTTTGACCAAGGATFCCATTTCTTCTA 1860
Db 1789 TTATGCTGATTAATTTTAGACTTAAATTTTCTTTGACCAAGGATFCCATTTCTTCTA 1848
Qy 1861 TCTTGTCTTCACTGCTGAGATTCCTCTCTATCTCTGTTGATTTCTGTCAGTGAGGCTTG 1920
Db 1849 TCTTGTCTTCACTGCTGAGATTCCTCTCTATCTCTGTTGATTTCTGTCAGTGAGGCTTG 1908
Qy 1921 TCTTGTGAGTTTCTGTTGGTCTTCTTAAATTTTCTTTTCCAGATTTCCCTCAGTTTGGGT 1980
Db 1909 TCTTGTGAGTTTCTGTTGGTCTTCTTAAATTTTCTTTTCCAGATTTCCCTCAGTTTGGGT 1968
Qy 1981 TTGTGTTTAAATTTCTATTTCCACTTTTCCAGTCTGAAATGTTTACTCATTTTCCCTCCC 2040
Db 1969 TTGTGTTTAAATTTCTATTTCCACTTTTCCAGTCTGAAATGTTTACTCATTTTCCCTCCC 2028
Qy 2041 AGTATTTACATTTTCTATAGGTTTCTTAAATGGAATTTATTTCAATTTCTTCAAGGACCTT 2100
Db 2029 AGTATTTACATTTTCTATAGGTTTCTTAAATGGAATTTATTTCAATTTCTTCAAGGACCTT 2088
Qy 2101 TTATGAATTCATAAATGATATGTTAAGTCTTGTGCTTGTGCTTGTGCTTGTGCTGCTG 2160
Db 2089 TTATGAATTCATAAATGATATGTTAAGTCTTGTGCTTGTGCTTGTGCTTGTGCTGCTG 2148
Qy 2161 TCAGGCTTATGTAATAGGTTTATAGGAGGACATATTTGCTGCTTGTGCTTGTGCTGCTG 2220
Db 2149 TCAGGCTTATGTAATAGGTTTATAGGAGGACATATTTGCTGCTTGTGCTTGTGCTGCTG 2208
Qy 2221 GTTTTGTCTTTGGCATAAGACGCTGAGTTTGGGATGATTTTAAATTTAGGTCCTGAT 2279
Db 2209 GTTTTGTCTTTGGCATAAGACGCTGAGTTTGGGATGATTTTAAATTTAGGTCCTGAT 2267

RESULT 4

US-10-664-775-3
; Sequence 3, Application US/10664775
; Publication No. US20040115776A1
; GENERAL INFORMATION:
; APPLICANT: Simesen, Ruth B
; APPLICANT: Pedersen, Anette A
; APPLICANT: Falsst, Steffen
; APPLICANT: Jensen, Jan J
; APPLICANT: Weillguny, Dietmar
; TITLE OF INVENTION: Method for Making Recombinant Proteins
; FILE REFERENCE: 6448.200-US
; CURRENT APPLICATION NUMBER: US/10/664,775
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: Danish Application PA 2002 01384
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 60/416,566
; PRIOR FILING DATE: 2002-10-07

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Cricetus griseus
; US-10-664-775-3

Query Match 87.9%; Score 2003; DB 1; Length 2003;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 2003; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 AGCTTTCCAGAGAGACTTTCATATAATATTTTCTTGAAGCTCTGCTGGCAATACTTCTG 166
Db 1 AGCTTTCCAGAGAGACTTTCATATAATATTTTCTTGAAGCTCTGCTGGCAATACTTCTG 60
Qy 167 GGCTGCTGCTTTCTCCCTGCTGATTTCTAGGTTAGGGTTACCACTGCTCTCTCTC 226
Db 61 GGCTGCTGCTTTCTCCCTGCTGATTTCTAGGTTAGGGTTACCACTGCTCTCTCTC 120
Qy 227 CTTTCTCTAACAATCTTGGCCAGAGTAGGGCACTACCGCAATCCCTCTCTCTCCAA 286
Db 121 CTTTCTCTAACAATCTTGGCCAGAGTAGGGCACTACCGCAATCCCTCTCTCTCCAA 180
Qy 287 ACATTTCTAATTTCTGATTTCTATCTTGGCTCAATTTTAACTCAGTAGTGGTTGG 346
Db 181 ACATTTCTAATTTCTGATTTCTATCTTGGCTCAATTTTAACTCAGTAGTGGTTGG 240
Qy 347 TTTCCATAGTTTGTAAAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 406
Db 241 TTTCCATAGTTTGTAAAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 300
Qy 407 GCTGTGCTGATCAGATAGGACATAGAGTATTTTCAATTTGCTTTTATCTGTCGAGACT 466
Db 301 GCTGTGCTGATCAGATAGGACATAGAGTATTTTCAATTTGCTTTTATCTGTCGAGACT 360
Qy 467 TGTGTTGTTTGAATATGATTTTCAATTTTGGAGTTTCAATAGGTTGCTGCAAGAAG 526
Db 361 TGTGTTGTTTGAATATGATTTTCAATTTTGGAGTTTCAATAGGTTGCTGCAAGAAG 420
Qy 527 TACAGTCTTGTCTTTTGGTGAATAGTCTGTAATATCTCTAGGTCCTCTGTTTATG 586
Db 421 TACAGTCTTGTCTTTTGGTGAATAGTCTGTAATATCTCTAGGTCCTCTGTTTATG 480
Qy 587 ACATCAGTTAGTCTCCAGCAATTTCTGTTTCTGTTTGTGAGATGACCTAACTGTTG 646
Db 481 ACATCAGTTAGTCTCCAGCAATTTCTGTTTCTGTTTGTGAGATGACCTAACTGTTG 540
Qy 647 AGAATGGGTTATGAGTGGCCACTATCTGTTGAGGTTCAATATGTTGTTTATG 706
Db 541 AGAATGGGTTATGAGTGGCCACTATCTGTTGAGGTTCAATATGTTGTTTATG 600
Qy 707 TGTAGCTGTCTTGTGTTTATGAATTTGGGTGACATTTGTTTGGTGCATAGACATTAAGA 766
Db 601 TGTAGCTGTCTTGTGTTTATGAATTTGGGTGACATTTGTTTGGTGCATAGACATTAAGA 660
Qy 767 ATTGCAATGCTCTTGGTGGATTTTCTTGTGATGCTATGTTGTTTCTTCCCAATCTC 826
Db 661 ATTGCAATGCTCTTGGTGGATTTTCTTGTGATGCTATGTTGTTTCTTCCCAATCTC 720
Qy 827 ATCTGCTTGTGTTTGGTGTGTTTAACTCTATAGTGCAGATATTTAAATGACATGTTGCTG 886
Db 721 ATCTGCTTGTGTTTGGTGTGTTTAACTCTATAGTGCAGATATTTAAATGACATGTTGCTG 780
Qy 887 CTTCTTAGGCCATTTGCTTGAATATCTTTTCCATCTTCTTACTCTAAGGTGATGCTA 946
Db 781 CTTCTTAGGCCATTTGCTTGAATATCTTTTCCATCTTCTTACTCTAAGGTGATGCTA 840
Qy 947 TCCATGGTAGGTTGCTTTTGGATGCGACAGTGGATGTTGTTTCTATATCCAT 1006
Db 841 TCCATGGTAGGTTGCTTTTGGATGCGACAGTGGATGTTGTTTCTATATCCAT 900
Qy 1007 TCTGTTACCCAGTATCTTTTCTTAGAGAAATTAAGATCATTTGATGATGTTGAGAA 1066

Db	901	TCTGTTACCCAGTATCTTTTCTAGAGAAATTAGATCATTGAGTCATTGATGTTGAGAA	960
Qy	1067	TTATCAATGAGCAGGTTTGGGATCTTGTTATCTTGCACTTGTGAAGTGTGTGTGT	1126
Db	961	TTATCAATGAGCAGGTTTGGGATCTTGTTATCTTGCACTTGTGAAGTGTGTGTGT	1020
Qy	1127	GT	1186
Db	1021	GT	1080
Qy	1187	GT	1246
Db	1081	GT	1140
Qy	1247	TTTCAATTTTCTTGAATCTGGTAAACATCTTTAGATTGAAGTTTTCCTTAGCCTTCTT	1306
Db	1141	TTTCAATTTTCTTGAATCTGGTAAACATCTTTAGATTGAAGTTTTCCTTAGCCTTCTT	1200
Qy	1307	TAGGTCTGCATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATGCTTTCT	1366
Db	1201	TAGGTCTGCATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAATGCTTTCT	1260
Qy	1367	TTCTCCAACTATTTGACAGAAAGTTTCTTAACTGAGTAGTCTGGCTGACATCTGTA	1426
Db	1261	TTCTCCAACTATTTGACAGAAAGTTTCTTAACTGAGTAGTCTGGCTGACATCTGTA	1320
Qy	1427	GTCTCTTGGAGTCTGTAGCACATCTGTGACGGGCTTCTTACATTTTGAAGTTTCTATTGG	1486
Db	1321	GTCTCTTGGAGTCTGTAGCACATCTGTGACGGGCTTCTTACATTTTGAAGTTTCTATTGG	1380
Qy	1487	AAAAGTCAGGTGTAATCTTAATACATCTGCGCTTATATGTTAATGTCCTTTTCCCTT	1546
Db	1381	AAAAGTCAGGTGTAATCTTAATACATCTGCGCTTATATGTTAATGTCCTTTTCCCTT	1440
Qy	1547	GCATCTTTTAAATTTCTTTCTTTGTTCTATCTTTTGTAGTGAATTTGATTAATGACACTGT	1606
Db	1441	GCATCTTTTAAATTTCTTTCTTTGTTCTATCTTTTGTAGTGAATTTGATTAATGACACTGT	1500
Qy	1607	GGGAGTCTTTTCCGGTCCAACTATTTGGGTGTTTGTATGCTTTGTACCTTGTATGACACTGT	1666
Db	1501	GGGAGTCTTTTCCGGTCCAACTATTTGGGTGTTTGTATGCTTTGTACCTTGTATGACACTGT	1560
Qy	1667	GGCATCTCTTTTCCAAAGGTAGGAAATTTTCTTTTGGTTTCTTGGAAATATTTTCC	1726
Db	1561	GGCATCTCTTTTCCAAAGGTAGGAAATTTTCTTTTGGTTTCTTGGAAATATTTTCC	1620
Qy	1727	CTGCTTTTGGACGTGCTTCTTCCCTTCTCTATTTCTTCTTGTGTTTGTGCAATGCTCT	1786
Db	1621	CTGCTTTTGGACGTGCTTCTTCCCTTCTCTATTTCTTCTTGTGTTTGTGCAATGCTCT	1680
Qy	1787	GGCTTCTTGGATGTTTATGCTGGAATTTTATGACTTACATTTTCTTGGACCAAGGT	1846
Db	1681	GGCTTCTTGGATGTTTATGCTGGAATTTTATGACTTACATTTTCTTGGACCAAGGT	1740
Qy	1847	ATCCATTTTCTTCTATTTGTTCTTCACTGCTGAGATTTCTCTTCTATCTCTTCTTATTTCT	1906
Db	1741	ATCCATTTTCTTCTATTTGTTCTTCACTGCTGAGATTTCTCTTCTATCTCTTCTTATTTCT	1800
Qy	1907	GTCACTGAGGCTGTCTGAGGTTCTGTTGGTTCTTAAATTTTCTTCAATTTCCAGATTT	1966
Db	1801	GTCACTGAGGCTGTCTGAGGTTCTGTTGGTTCTTAAATTTTCTTCAATTTCCAGATTT	1860
Qy	1967	CTTTCAGTTTGGGTTTGTGTTTATTAATTTCTATTTCCACTTTTCAGGCTCTGAAATGTTTAA	2026
Db	1861	CTTTCAGTTTGGGTTTGTGTTTATTAATTTCTATTTCCACTTTTCAGGCTCTGAAATGTTTAA	1920
Qy	2027	CTCATTTTCCCTCCAGTATTTACATTTTCAATGTTTCTTTTAAATGGAATTTATTCATTTCC	2086
Db	1921	CTCATTTTCCCTCCAGTATTTACATTTTCAATGTTTCTTTTAAATGGAATTTATTCATTTCC	1980
Qy	2087	TCTTCAAGGACCTTTTATGAATT	2109

Db	1981	TCTTCAAGGACCTTTTATGAATT	2003
RESULT 5			
US-10-087-192-577			
; Sequence 577, Application US/10087192			
; Publication No. US20020182586A1			
; GENERAL INFORMATION:			
; APPLICANT: Morris, David W.			
; APPLICANT: Engelhard, Eric K.			
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR			
; TITLE OF INVENTION: CANCER			
; FILE REFERENCE: 529452000122			
; CURRENT APPLICATION NUMBER: US/10/087,192			
; CURRENT FILING DATE: 2002-03-01			
; PRIOR FILING DATE: 2000-12-22			
; PRIOR APPLICATION NUMBER: US 09/747,377			
; PRIOR FILING DATE: 2000-12-22			
; PRIOR APPLICATION NUMBER: US 09/798,586			
; PRIOR FILING DATE: 2001-03-02			
; NUMBER OF SEQ ID NOS: 2059			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 577			
; LENGTH: 41219			
; TYPE: DNA			
; ORGANISM: Mus musculus			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: (1)...(41219)			
; OTHER INFORMATION: n = A,T,C or G			
US-10-087-192-577			
Query Match 27.5%; Score 626.9; DB 1; Length 41219;			
Best Local Similarity 65.8%; Pred. No. 1.8; Gaps 23;			
Matches 1223; Conservative 0; Mismatches 476; Indels 161;			
Qy	292	TCTATTCTTGTATTTCTATCTTTGGCTCAATTTTAACTCAGTAGTAGTGTGTTTGGTTTCC	351
Db	37261	TTTATTTCTTTATTTCTTCCCTGACTCAGAGATCATTTGAGAAGGGTGTTCATTTTCC	37320
Qy	352	ATAAGTTTGTAGTATTCTGTTTCTGTTTCTGTTTGTGTTTATCTAGATTAGCTGT	411
Db	37321	ATGAATTTGTAGGCTTTCTGTTTCTGCTGTTGTGAAG--TCCAGCTTTTATGCCAT	37377
Qy	412	GGTGGTCAGATAGGACATAGAGTATTATTTCAATTTGTTTATCTGTCGAGACTTGCTT	471
Db	37378	GATGGTCTGNTAAGATACAGGCAATTTTCAATTTCTTGTATCTGTTGAGGCTTGCTT	37437
Qy	472	TGTTTGAATATGTAATTCATTTTGGAGA--GTTTCAATAGGTTCTGACAGAGGTA-	528
Db	37438	TGGACCAACTATATGTTTGGTGAAGTAGTCTGTAATATCT-CTAGTCCACTTGGTTAT	585
Qy	529	--CAGTCTTTGTTTGGTGAAGTAGTCTGTAATATCT-CTAGTCCACTTGGTTAT	585
Db	37498	TTTTTTTTTATGTTTGGTGAAGTATTCATAGGTGTCATATAGTCCATTTTCAATTCAT	37557
Qy	586	GACATCAGTTAGTCCAGCAATTTCTGTTTCTGTTTCTGTTTGTGAGATGACCTPAATGTTG	645
Db	37558	AATGTCTACTAGTTTCTTTTATTTCTGTTTAG-TTCTGCTCTGATGACATGTCAATTG	37616
Qy	646	GAGAGATGGGTATGGAAGTAGCCCACTATCTGTTGTGAGG-TCAATATGTTGATTTTA	704
Db	37617	GTGAGTAGGTTTCTTTTATGAGTGTGGTGGCTTGTGTTTGGGTCATAGATTTTCA	37671
Qy	705	GCTGTAGCTGTGCTGTTTATGAACTTGGGTGACATTTGTGTTTGTGTCATAGACATTA	764
Db	37672	GCTTTAGCAGTGTTCCTTTTATGAGTGTGGTGGCTTGTGTTTGGGTCATAGATTTTCA	37731
Qy	765	GAATTCGAATGCTCTTGGTGA-TTTTCCCTTTGATGCCATATGATTTTCTTCCCAAT	823
Db	37732	GAATTCGAATGCTCTTGGTGA-TTTTCCCTTTGATGCCATATGATTTTCTTCTTACGT	37791
Qy	824	CTCATCTGCTTAGTTTGGGTTTAAAGTCTATTAGTCAGATATTAATAATGACTGTATCGGC	883

Db	37792	CTCTTTTGATTAATTTTGGTTGAAGCTTATATCTTTAGAAATAGAAATGGTTACTCGACG	37855
Qy	884	TTGCTCTTTAGGGCCCAATTTGCTTTAGAAATA-TCTTTTCCATCCTCTTTATCTCTAAAGTGATG	942
Db	37852	TTGCTCTTTTGGTCTGTTTGGTTGGAAAACCTTTTTTCCAAACCTTTTCTCTGAGATAATG	37911
Qy	943	TCCTA---TCCATCGGTAGTGTCTTTTGGATGACAGCAGTAGGATGATCTGTGTTTCA	999
Db	37912	TCATATTTTATTCCTGAGTGTGTCTTGTGTACACAGAAATAAGATGATCCTGTGTTTCA	37971
Qy	1000	TATCCATTCCTGTTACCACAGTATCTTTTCTAGAGAAAATTAAGATCAATGAGTCATTGATG	1059
Db	37972	CAITCATCTGTTAGCTGTGCTTTTATTTGGGGAATTGAG-----TCCACTGGTG	38023
Qy	1060	TTGAGAAATTAACAATGAGCAGTGTTTGTGGATTCCTTGTTATCTTGCACTGTGGAAGTG	1119
Db	38024	TTGAGAGATAATAATGACCAGTGAAATTTATTTTCAGGTATTTTGTCTGTGTAGTGTATC	38083
Qy	1120	TGT	1179
Db	38084	TGT	38117
Qy	1180	TGT	1239
Db	38118	-----CTTTGCTGTCTTTTCTGGTATGAGTTA	38146
Qy	1240	TTTATTATTACATATTTTCTGAATGGGTAACTTTTAGATTCAAGTTTCTCCTAG	1299
Db	38147	CTTAGTTACTATGTTTCTTGGATGAGTTATCTTCTGGAGTTGGAATTTTCTCTCTAG	38206
Qy	1300	CCTCTCTT--TAGGTCTGCATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAA	1357
Db	38207	AAATTTTGGTAGGGCTGCATTTGTGGCTAGATATTGCTTAAATTTGGAT-----	38255
Qy	1358	TGTCTTCTTTCTCCAACTATTTGTGTACAGAAAGTTTCTAAGTCAGTAGTCTGGCCTG	1417
Db	38256	-----CTATGGTAATTGAGAGCTTGTCTGAGTATGTCTGAGTATGTAGTCTAGGTTG	38298
Qy	1418	ACATCTGTAGTCTCTTTGGA-GTCTGTAGCACATCTGTGCAGGGCTTCTTACATTTTGTAG	1476
Db	38299	GCACTGTGTGTTTTTAGAGTTTCAAGATAATCTGTCCAGTCTCTCTAGCTTTTAGAG	38358
Qy	1477	TTTCTATTGAAAAGTCAGGTGTAATCTTAATACATCTGTCCCTTTATATGTTAAATGGTCT	1536
Db	38359	TCCTCTGTT--GAGAAATTTGAATGTAAATCTGATAGGTGACGCTTTATATGTTACCT-GTCC	38416
Qy	1537	TTTTTCCCTTGGCATCTTTTAAATCTTCTTTTGTCTATACTTTTAGTGAATTTGATTTAT	1596
Db	38417	TCCTTCCCTCACAGCTTTTAATCTCTTTTCTTCTTAAGATTGTGTGTTTGTATTTAT	38476
Qy	1597	TATGCATGTGGGAGTTCTTTTCCGGTCCAAATCTAATTGGTGTGTTTGTATGCTCTTG	1656
Db	38477	TATATAGCAGAGAAATTTCTTTTCTGTGCCAGTCTA-ATGGTGTCTTATAAGCTTTTGTG	38535
Qy	1657	TACCTTGATAGGCATCTCTTTCTCAGGTTAGGAATTTTCTTTTGTGTTTCTTTGAA	1716
Db	38536	TATGTTATAGGCAACTCTTTCTTTAGGTTGGAAAGTTTCTCTATGATTTTGTGGAG	38595
Qy	1717	AAATATTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCCCTCTATTC---1763	
Db	38596	AGTATTTCTGGGCCCTTTGGAGCTGGGACCTTTCACTTCTCTCATTACGTTATCTTAG	38655
Qy	1764	-TTTGGTTTTGATAGTAGTCTGTGGCTTCTGTGATGTTTTTAGCCTCGAATATTTAGA	1822
Db	38656	GTTAGTCTTTTCATGGTATCCACAGATTTCTGTGATGTTTTGTGTGAGAAATTTTAGT	38715
Qy	1823	CTTAAACATTTCTTTGACCAGGATATCCATTTCTTCTATCTGTGCTTCACTGCTCGAGAT	1882
Db	38716	TTTAAACATTTCTTTGACTGGTGA-----TCATCTA-TGCCCTAGAT	38757
Qy	1883	CTCTCTCTCTATCTCTGTGATATCTGTCTAGTAGGCTGTGCTGTAGGTTCTCTGTGGGTT	1942
Db	38758	CCTCTCTTCCATCTCCTGTCTTCTGTGTGTGATGCTGTGCTCTGTGTGTTCTGTGTCTT	38817

```

1943 CTTAAATTTTTCATTTCCAGATTTCCCTTCAGTTGGGGTTTGTATTAATCTATTTCC 2002
Db 38818 CCCTAGTTTTCACCTCCAG-----GGAAGTTTATGTTTCTTTTATTACTTCTATTTCC 38871
Qy 2003 ACITTCAGGTCCTGAAATGTTTTACTCATTT--TTCCCTCCAGATTTTACATTTTCATAG 2059
Db 38872 ACITTCAGGTCCTACACAGTTTATTATTTCCCTTCATCTGTTTAAATATTATTTCTTAT 38931
Qy 2060 GTTTCCTTTAATGAAATTTATTCATTTCCCTCTTCAAGGACCTTTTATGAATTCATAAAATGT 2119
Db 38932 ATATCTTTTAGGGATTAATTTGTTTCCCTCTTTAAAGACATCTACATGTTTGATTTGTTATTT 38991

RESULT 6
US-10-087-192-1663
; Sequence 1663, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1663
; LENGTH: 193853
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(193853)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1663

```

Query Match	26.8%;	Score 510.3;	DB 1;	Length 193853;
Best Local Similarity	63.8%;	Prod. No. 1.1;		
Matches 1180;	Conservative	0;	Mismatches 552;	Indels 117; Gaps 18;
QY	292	TCTATTCTTGATTTCTATCTTGCGCTCATTTTAACTACAGTAGTCAGTGTGTGGTTTCC	351	
Db	152240	TTTCTTTCTTTATTCCTTCCTTGACCAAGGTATCATTTGAGAAGAGTGTGTTCAGTTTCC	152299	
QY	352	ATAAGTTTGTAAAGTTTCTGTGTGTTTCTGTGTTTGTGTTTATCTAGATTAAAGCTGT	411	
Db	152300	ACGTGAATGTGGCTTCTCTATTTATTTTGTATTGAAGAT--CAGCCTTAGTCCAT	152356	
QY	412	GGTGGTCAGATAGACATAGAGTATTATTTCAATTGCTTTTATCTGTCGAGACTTGCTT	471	
Db	152357	GGTCATCTGATAGGATGATGGAACAATTTCAATGTTTGTGATATGTGAGGCTGTGTT	152416	
QY	472	TGTTTTCGAAATATGATTTCAAATTTGGAGA--GTTTTCATAGGGTGCTGACAAGAAGGTAC	529	
Db	152417	TGTGACCAATTATGTGGTCAAATTTGGAGAAGGTACCATGAGGTGCTGAGAAGAAGGTAT	152476	
QY	530	AGTC-TTTGTGTTTTGGTGAATAGTCGTGAAATATCTCT-AGTCGCACTTGTTTATGA	587	
Db	152477	ATCCTTTTGTTTTAGGATAAAATGTTCTGTAGATATCTGCAGATCCATTTGTTTCAATA	152536	
QY	588	CATCAGTTTAGTCCAGCATTTCTCTGTTTTCGTTTTTGTTCGAGATGACCTAACTGTTGGA	647	
Db	152537	CTTCTGTAGTTTCACTGTGTCCCTGTTTAG-TTTTTGTTTCCAGATCTGTCACATAGT	152595	
QY	648	GAGAATGGGTTTGAAGTAGTACGCCCACTATCTGTGTGTAGGT-CAATATGTGATTTTTAGC	706	
Db	152596	GAAAGTGGTGTGTTGAAGTCTCCCACTATTATTGTGTAGGTGAAGTGTGTCGCTTCAGC	152655	

QY 707 TGTAGCTGTCTGTTTATGAACCTGGGTGACATTTGTTTGGTGATAGACATTAAGA 766
 Db 152656 TTACTAAAGTTCTTTAATGAATGTGGCTGCCCTGTTATTTGGAGCATAGAAATT CAGA 152715
 QY 767 ATTGCAATGCTCTCTGGTGTGATTTT-CCITTTGATGCTGTAGTATTCTTCCCATCT 825
 Db 152716 ATTGATAGTTCTCTTGGAGATTTTACCTTTTGAATATGAAGTGGCCCTCCTCTGCT 152775
 QY 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATTAATAATGACTGTATCGGCT 884
 Db 152776 TTTTGTATGACTTTGGGTTGGAGTCAATTTTATTAGATATTAGATGGCTACTCCAGCT 152835
 QY 885 TGCTTCTTAGGCCATTTGCTTAGAATA-TCTTTTCCATCCTTTTACTCTAAGTGATGT 943
 Db 152836 TGTTTTTCATACTATTGCTTGGAAATTTGTTTCCAGCCCTTTCAITCTGAGGTAGTGT 152895
 QY 944 CTATCCAT--GGTAGGTGTCTTTTGGATGACAGTAGTAGGATGATCTGTGTTTCAT 1000
 Db 152896 CTAICTTTTCTCTGAGATGAGTTTCTCTAAGCAGCAAAATGTTGGTCTTGTGTTGT 152955
 QY 1001 ATCCATCTGTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCATGATGT 1060
 Db 152956 AGCCAGTTAGTTAGTCTATGCTGTTTATTGGGAGTTGAG-----TCCATTGATAT 153007
 QY 1061 TGAGAAATTATCAATGACGAGTGTGTTGATCTTGTATCTTGCACCTGTGAGTGTGT 1120
 Db 153008 TAAGAGATATTAAAGAAAGTAATGTTGCTCTCTGTTATTTTCTTTTCTTTTAAAG 153067
 QY 1121 GT 1180
 Db 153068 TTGGCAATCTGTCTTGTGGCTGCTCTTCTTTTAGTTTG----- 153106
 QY 1181 GT 1240
 Db 153107 -----TTGAGGGAT 153115
 QY 1241 TTATTATTCAATTTTCTTGAATGTGGTAAATCTTTTAGATTGAAGTTTTCTCCTAGC 1300
 Db 153116 TACCTTCTGTCTTTTCTAGGGCGTGGTTCGGTCTGTATGTTGTTTCTATTATT 153175
 QY 1301 CTTCTTT--AGGTCTGCAATTTGAAGATAGATATTTCTTACATCTGATTTTATCTTGAAT 1358
 Db 153176 ATCCCTTGAAGGGCTGGAATTTGTGAAAGATAATGTGTGAATTTGGTTTGTCTGTGAAT 153235
 QY 1359 GTCTTCTTCTCCCACTATTGTGACAGAAATTTTCTAAGTCAGTAGTCGGCTGA 1418
 Db 153236 ACTTTGGTTTCTCCATCTATGTAATTCAGAGATTGGCTGGGTATAGTAGCTGGCTGG 153295
 QY 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGACGGCCCTTCTTACATTTTGAGTT 1478
 Db 153296 AATTGTGTCTCTTTAGTGTCTGTATAACATCTGTCCAGGCTCTTCTGGCTTTTATAGTC 153355
 QY 1479 TCTATTGAAAAGTCAGGTGAATTTCTAATACATCTGCTTTTATATGTTAATGCTCTTT 1538
 Db 153356 TCTGCT-GAAAAAATCTGTAATTCGATAGGCTTGGCTTTATATGTTACTT-GACCTT 153413
 QY 1539 TTTCCCTTGCATCTTTTAAATATCTTCTTGTGTTCTATATCTTTTAGTGATTGATTATTA 1598
 Db 153414 TTTCCCTTACTGCTTTTAGTATTCTATCTTTATTAGTGCAATTTGTTCTGTGATTATTA 153473
 QY 1599 TGACGTGGGAGTTTCTTTTCCGGTCCAAATCTATTATTTGGTGTGTTTGTATGCTCTTGTA 1658
 Db 153474 TGTGTGGGAGGAATTTCTTTTCTGCTCCAGTCTATTTTGGAGTTCTGTAGGCTTCTTGT 153533
 QY 1659 CTTGATAGGCAATCTTTCTCAAGGTTAGGAATTTTCTTTTGTGTTTCTTGAAAA 1718
 Db 153534 TGTTCATAGGCAATCTTTTCTTTTAGATTGGAGTTTCTTCAATATTTTGTGTGAGA 153593
 QY 1719 TATTTCCTGCTTTTGGACCTG-----CCTTCTTCCCTTCTCTATTCTCT 1764
 Db 153594 TATTTCCTGCTTCTTTGAGTTGAAATCTTCTTCTTCACTCTCTTATTATCTGTAGGT 153653

QY 1765 TTGGTTTTTGCATAGTGTCTCTGGCTTCTGATGTTTTTATGCCCTGGATTTATTAGACT 1824
 Db 153654 TTGGTCTTCTCATTGTGCTCTGGATTTCTCTGGATTTTGTAGTTAGGATCTTTTTCGCA 153713
 QY 1825 TAACATTTTCTTTGACCAAGGTATCCATTTCTTATCTTGTCTTCACTGCTGCTGAGATTC 1884
 Db 153714 TTCCATTTTCTTTGATTTGTTGCGCATGTTCTCTATGGAATCTTCTGACCTGAGATTC 153773
 QY 1885 TCTCTTCTATCTCTGTATTTCTGTAGTGAGGCTGTCTCTGAGTTTCTG-TTGGGTTT 1943
 Db 153774 TCTCTTCCATCTCTGTATTTCTGTGTGCTGATCTCAAAATCTGTGTTCCAGATTTCTTTC 153833
 QY 1944 TTAATTTTCTTCCAGATTTCTTCACTTGTGTTGTTTGTGTTTATTAATCTTATTTCCA 2003
 Db 153834 CTAGGTTTCTATCTCCAGCGTTGCTCACTTTGGGTTTCTTTATTTGTGCTACTTCTC 153893
 QY 2004 CTTTCAGGTCCTGAAATGTTTTTACTCATTTTCTCTCC---AGTATTTACATTTTCATAGG 2060
 Db 153894 TTTTGTAGGTCCTTGGATGTTTTTATTCATTTCTTCACTGTTGTTGTTGTTTTCATGCA 153953
 QY 2061 TTTCTTTAATGATTTATTCATTTCTCTCTTCAAGCACCTTTTATGAAT 2109
 Db 153954 ATTCTTTAAGGATTTTGTGCTTCTCTTTTAAATGTCTTCTACTTGT 154002

RESULT 7
 US-10-369-214-129/c
 ; Sequence 129, Application US/10369214
 ; Publication No. US20030232037A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Groot, Pieter C.
 ; APPLICANT: Berghenhegouwen van, Bram J.
 ; APPLICANT: Oosterhout van, Antoon J.M.
 ; TITLE OF INVENTION: Genes involved in immune related responses observed
 ; TITLE OF INVENTION: with asthma
 ; FILE REFERENCE: P53837US00
 ; CURRENT APPLICATION NUMBER: US/10/369,214
 ; CURRENT FILING DATE: 2003-02-15
 ; PRIOR APPLICATION NUMBER: EP 00202867.8
 ; PRIOR FILING DATE: 2000-08-16
 ; PRIOR APPLICATION NUMBER: PCT/NL01/00610
 ; PRIOR FILING DATE: 2001-08-16
 ; NUMBER OF SEQ ID NOS: 139
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 129
 ; LENGTH: 19619
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)..(19619)
 ; OTHER INFORMATION: /note="Contig 1A gene"
 US-10-369-214-129

Query Match 26.0%; Score 591.499; DB 1; Length 19619;
 Best Local Similarity 63.2%; Pred. No. 3.5; Indels 119; Gaps 16;
 Matches 1138; Conservative 0; Mismatches 545;

QY 292 TCTATTCTTGTATTTCTATCTATCTGGCTCAATTTTAACTCAGTAGTGTGTTTGGTTTCC 351
 Db 19360 TTTCTTTCTTTTCTCTTCTTCCCTTGACCAAGGTATCATTTGAGAAGAGTGTGTTGTTCC 19301
 QY 352 ATAGTTTGTAGTTTCTGTTGTTTCTGTTGTTGTTGTTTCTAGATTTCTAGATTTAGCTGT 411
 Db 19300 ATGTGAATGTGGCTTCTGTTATTATTGTTTATTGAAT---CAGCCTTAGTGAT 19244
 QY 412 GGTGGTTCAGATAGGACATAGAGTATTATTCAATTTGCTTTTATCTGTCGAGACTTGTCT 471
 Db 19243 GGTGATCTGATAGGATACATGGACAAATTTCAATATTTTGAACTCTGTTGAGCCTGAT 19184
 QY 472 TGTTTTGAATATGTTATTCATTTTGGAGA--GTTTCTAGGGTCTGTACAGAGAGGTAC 529
 Db 19183 TGTGACCTATTATGTGGTCAATTTTGGAGAGGTACCATGAGTGTCTGAGAGAGAGGTAT 19124

Qy 530 AGTC-TTGTGTTTGTGGAATAAGTCTGTAAATATCTCT-AGTCCACCTGGTTTATGA 587
Db 19123 ATCCTTTTGTGTTTGTAGGGTAAATGTTCTGTAGATATCTGTAGATCCATTTGTTTCATCA 19064
Qy 588 CATCAGTTAGCTCCAGCAATTTCTGTTTCCGTTTGTGTTTGTGAGATGACCTAACTGTGTGA 647
Db 19063 CTTCCTGTTAGTTTCTAGTGTGCTCTGTTTGTAG-TTTCGTGTTTCCATGATCTGTCCATTGGT 19005
Qy 648 GAGAATGGGTATTTGAAGTAGCCCACTATCTGTGTGTGAGG-TCAATATGTGATTTTATG 706
Db 19004 GAAAGTGTGTGTGGAAGTCTCCCACTATTTATTTGTGTGAGCGCAATGTGTGCTTTTGAGC 18945
Qy 707 TGTAGCTGTGCTGTTTATGAACTCTGGTGACATTTGTTGTTGTGTGATGACATTAAGA 766
Db 18944 TTTACTAAGTTTCTTTAGTGAATGTGGCTCTCTTTGTTTGTGAGATAGATATTGAGA 18885
Qy 767 ATTGCAATGTCCTCTGTGTGATTTT-CCTTTGTAGTCTATGTAGTATCTTCCCAATCT 825
Db 18884 ATTGAGAGTTCTCTTGTGAGATTTTACCTTTGTGATGAGATGAAGTGTCCCTCTGTCT 18825
Qy 826 CATCTGCTTAGTTTGGTTTAACTATTAAGTCTATTAAGTCTAGATATTAATAAGTCTATCGGCTT 885
Db 18824 TTTTGTAGTACTTTGGTTGGAAGTCTATCTATCAGATATTAAGATGGCTACTCTGCTT 18765
Qy 886 GCTTCTTAGGGCCATTTGCTTAGAATATCTTTTCCATCCTTTTACTCTTAAGGTGATGCT 945
Db 18764 GTTCTCTCATACCAATTTGCTTGGAAATGTTTCCAGCCTTTCACTCTGAGGTAGTGTCT 18705
Qy 946 ATCCAT---GGTAGTGTCTTTTGTGATGAGCAGTAGGATGATCTTTGTTTCATAT 1002
Db 18704 ATCTTTTCTCTGAGATGAGTTTCTCTGTAAGCAGCAAAATGTTGGTCTTTGTTGTAG 18645
Qy 1003 CCATCTGTATCCAGATATCTTTTCTAGAGAAATTAAGATCATTTGATCATTTGATGTTG 1062
Db 18644 CCAGT-TGTTAGTCTATCTTTTATTTGGGAGTTGAGA-----CCAATTGATGTTA 18594
Qy 1063 AGAATTAATCAATGAGCAGTGTGTTGGATCTTGTATCTTGCACCTTGTGAAGTGTGT 1122
Db 18593 AGAGATTAAGAAAGTAATCTGCTTCCCTGTTATTTAGTTGTTAAAGTGGCAAT 18534
Qy 1123 GT 1182
Db 18533 CTGTTCTGTGGCTGTCTCTTTTAGGTTG----- 18503
Qy 1183 GTCTGTGTGTGTGTGTGTGTCTCTCCCTCTTTTGTATTTTGGCCCTGGAATTTATTT 1242
Db 18502 -----TTGAGGGAATTA 18492
Qy 1243 ATTATTCATATTTTCTTGAATGTGGTAACATCTTTAGATTGAAGTTTCTCTAGCCCT 1302
Db 18491 CCTCTGTTTCTTAGGGCTGTTGTTCCCGTCTCTGTATTTGGTTTCTGTTATTTAT 18432
Qy 1303 TCTTTT-AGGTCTGCATTTGAGATAGATATCTTTACATCTGATTTTATCTTAGATGT 1360
Db 18431 CCTTTGAAGGCTGATTCGTGGAGAGATAATCGTGAAATTTGGTTTGTGCGTGAATC 18372
Qy 1361 CTCTCTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCCTGACA 1420
Db 18371 TTTGGTTTCTCCTCTATGATAATTGAGAGTTTGGCTGGGTATAGTAGCCTGGGCTGAG 18312
Qy 1421 TCTGTAGTCTCTTGAGTCTCTAGACATCTGTGCGAGGCTCTCTTACATTTTGTAGTTTC 1480
Db 18311 TTTGTGTCTCTTAGTGTCTGTATAACATCTGTCCAGGCTCTCTCTGGCTTTTCATAGTCTC 18252
Qy 1481 TATTGAAAAGCTAGGTTGATTTCTAATACATCTGCTCTTATATGTTAAATTTGGTCTTTT 1540
Db 18251 TGGT-GAAAAATCTGTTGTAATCTGATAGGCTTGCCTTTTATATGTTACTT-GACCTTTT 18194
Qy 1541 TCCCTGCACTTTTAAATTTCTTTCTTTGTTCTATATCTTTTGTAGTATTTGATTTATG 1600
Db 18193 TCCCTTACTGCTTTTAGTATTTCTATCTTTATTTAGTGAATTTGATTTCTGATTTATG 18134

Qy 1601 CACTGTGGGAGTCTCTTTTCCGGTCCAATCTATTGTTGGTGTGTTTGTATGCTTCTTGTACC 1660
Db 18133 TGTCCGGAGGAATTTCTTTTCTGGTCCAGTCTATTGAGAGTCTCTAGGCTTCTTGTATG 18074
Qy 1661 TTGATAGGATCTCTTTTCTCAAGGTTAGAAATTTTCTTTTGGTTTCTTGAATA 1720
Db 18073 TTCAATAGCACTCATCTTTTGAATTTGGAAGTCTTTCTCAATATTTTGTGAAGATG 18014
Qy 1721 TTTTCCCTCTCTTTTGACCTG-----CCTTCTCCCTTCCCTCTATTCCTTT 1766
Db 18013 TTTGTGGACCTTTGAGTTGAAATCTTCAATCTCATCCACTCTATTATCCGTAGCTTT 17954
Qy 1767 GGTTTTGCATAGTGTCTCTGGCTTCCCTGATGTTTATGCTCGATTTTGTAGCTTTA 1826
Db 17953 GGTCTTCTTATTTGTGCTGGAATTTCCGTATATTGTTGAGTAGGATCTTTTGTCAATTT 17894
Qy 1827 ACATTTTCTTTTGACCAAGGATCTCAATTTCTTCTATCTTCTTCTTCACTGCTCGAGATCTC 1886
Db 17893 CCATTTTCTTTGATTTGTCGCGATGTTCTATGGAATCTTCTGCACCTGAGATCTC 17834
Qy 1887 TCTTCTATCTCTGTATTTCTGTGAGGCTTGTCTGTAGGTTCTCTGAGGTTCTCTG-TTGGGTTCTT 1945
Db 17833 TCTTCCATCTCTGTATTTCTGTGTGATGCTCAAACTCTATGTTTCCAGATTTCTTTCT 17774
Qy 1946 AATTTTCTCAATTTCCAGATTTCTTCCAGTTTGGGTTTGTGTTTATTAATCTATTTCCACT 2005
Db 17773 AGGTTTCTATCTCTAGCGTTGCTTGCCTTGTGATTTCTTATTTGTTGTTCTACTTCCCT 17714
Qy 2006 TTCAGTCTGTAATGTTTCTTACTCAATTTTCTCCAGTATTTTACATTTTTCATAGGTTTCT 2065
Db 17713 TTTAGTCTAGTATGTTTGTGTTTCTTATTTTCCATCACCTGTTGTATGTTTCTTCTTCTT 17654
Qy 2066 TT 2067
Db 17653 TT 17652

RESULT 8

US-10-087-192-1945
; Sequence 1945, Application US/10087192
; Publication No. US20020192586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1945
; LENGTH: 178825
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(178825)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1945

Query Match 25.9%; Score 590.199; DB 1; Length 178825;
Best Local Similarity 63.4%; Pred. No. 1.3;
Matches 1297; Conservative 0; Mismatches 523; Indels 226; Gaps 28;

Qy 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGTGTTGTTCC 351
Db 98100 TTTCTTTTGTAAATTTCTGTCTTGACCAATCTATATTCAATAGTAGTGTGTTCAAGTTCC 98159

Qy	2004	CTTTCAAGTCTCGAAATGTTTTACTCATTTCTCCCAAGTAATTTACATTTTCATAGGTTT	2063
Db	946	TTTTTAATTCCTCAACATGTT-----TCGATTGTGTTTTCTCGGAATT	905
Qy	2064	CTTTAAATGGATTATTCATTTCCCTCTTCAAGACGCTTTATGAATTCATAAAATGTATGT	2123
Db	904	CTTTTCAGGATTTTGGCAATTCCTCTCTGTAGGCTTCTACTTGTTCCTAAGGGAGTTCT	845
Qy	2124	TAAAGTCTCTT 2133	
Db	844	TCACGTCTTT 835	

RESULT 11

US-10-388-934-790/c

; Sequence 790, Application US/10388934

; Publication No. US20040005547A1

; GENERAL INFORMATION:

; APPLICANT: Boess, Franziska

; APPLICANT: Suter-Dick, Laura

; APPLICANT: Wolf, Detlef

; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY

; FILE REFERENCE: 21199

; CURRENT APPLICATION NUMBER: US/10/388,934

; CURRENT FILING DATE: 2003-03-14

; PRIOR APPLICATION NUMBER: 02005336.9

; PRIOR FILING DATE: 2002-03-14

; PRIOR APPLICATION NUMBER: 02015657.6

; PRIOR FILING DATE: 2002-07-17

; NUMBER OF SEQ ID NOS: 862

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 790

; LENGTH: 6335

; TYPE: DNA

; ORGANISM: Rattus norvegicus (No. US20040005547A1way rat)

US-10-388-934-790

Query Match	25.8%	Score 588.5	DB 1	Length 6335
Best Local Similarity	62.7%	Pred. No. 5.8		
Matches 1173	Conservative 0	Mismatches 560	Indels 137	Gaps 18
Qy	292	TCTATTTCTTGATTTCTATCTTGCTCATTTTAACTCAAGTAGTGAGTTGTTGGTTTCC	351	
Db	2595	TTTCTTTCTTTATTTCTTCCCTTGACCAAGTTATCATTCAGTAGAGCATGTTTCAGTTTCC	2536	
Qy	352	ATAAGTTTGTAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTATCTAGATTTAAAGCTGT	411	
Db	2535	AAGTATATATGGGCATTCTCTCTGTAATTTATTGAAGACCAGTTTATAGGCCGT	2482	
Qy	412	GGTGGTCAGATAGGACATAGAGTATTATTCAAATGTCTTTTATCTGTGCAGACTTGCCT	471	
Db	2481	GGTGGTCGATAGCATGCAATGGGATTTTCTATCTTCTGACTGTTGAGGCCGCTTT	2422	
Qy	472	TGTTTTGAAATATGATTTTCAATTTTGGAG--AGTTTCATAGGGTGCTGACAAGAAGGTAC	529	
Db	2421	TTTGACCAAATATATATGTTGTCAAATTTTGGAGAAAGTACCATGAGGAGCTGAGAAGAAGTAT	2362	
Qy	530	AGTCCTTTGTGTTTTGG-TGAAATAGTCTGTAATAAT-CTCTAGGTCACACTTGGTTTATGA	587	
Db	2361	ATCCTTTTGGCTTTAGGATAGAAATGTTCTATATAAATATCCGTTAAGTCCCTTGGCTCATGA	2302	
Qy	588	CATCAGTTAGTCCAGCATTTCTCTGTTTCGTTTTGTTTGGATGACCTTAACCTGTGTGA	647	
Db	2301	CTTCTCTTAGTCTGCTACATCTCTGTTT-AAATTTCTGTTTCCCATGATCTGTCCATTGAT	2243	
Qy	648	GAGAACTGGGTTATGAAGTAGTACCCACATATCTGCTGTGAGGT-CAATATGTGATTTTATGC	706	
Db	2242	GAGAGTGGGGTGTTGAATCTCCACCATATATATGTGTGAGGTGCAATGTGTTTTCGAC	2183	
Qy	707	TGTAGCTGTGCTCTGTTTTATGAACCTCGGTCACATTTGTGTTGTCATAGACATTAAGA	766	
Db	2182	TTTAGTAAGGTTTCTTTTACATATGTTAGTCCCTTGTATTTGGGCATAGATATTTAGG	2123	

Qy	767	ATTGCAATGTCCTCTTGGTGA-TTTTCTTTTGAAGCCTATGTAGTATPCTTCCAAATCT	825
Db	2122	ATTGAGAGTTTCATCTCGTGGATTTTTCTTGTATGAATATGAAGTGTCTCTTCTTATCT	2063
Qy	826	CATCTGCTTACTTTTGGTTTAAAGTCTA-TTAGTCAGATATTAAATAGCACTGTATCGGT	884
Db	2062	TTTTTGAAGACTTTTAGTTAAAAATTCATTTTATTGATATTAGAAATGGCTACTCCAGCT	2003
Qy	885	TGCTTCTTAGGGCCATTTGCTTAGAATA-TCTTTTCCATCCTTTTATCTCTAAAGTGATGT	943
Db	2002	TGCTTCTTCTGACCATTTGCTTGGAAATTTGTTTCCAGCCTTTCACTCTGAGGTAATGT	1943
Qy	944	CTATCCATG--GTAGGTTCTTTTTTGGATGCAGCAGTAGGATGGATCTTGTTTTCAT	1000
Db	1842	CTGTCTTTGTCTCTGAGGTGTGTTTCTGTAGGCAGCAGAATGAGGGTCTCTGTCCGT	1883
Qy	1001	ATCCATTTCTGTATCCCAAGTATCTTTTTCTTAGAGAAATTAAGATCAATTGAGCTATTGATGT	1080
Db	1882	ATCCAGTTTGTAAATCTATGCTTTTTTATTTGGGAGTTAAG-----GCCATTTATGT	1831
Qy	1061	TGAGAAATATCAATGACAGCTGTTTGTGAATCTTGTATCTTGCACTTGTGAAGTGTGT	1120
Db	1830	TGAGAGATATTAAGGAATAGTGAATATGTCTCCGTTATATTCAATTTTGGATGTGAGG	1771
Qy	1121	GTCT	1180
Db	1770	TTATGTTTGTGTG-----	1758
Qy	1181	GTGTGTGTGTGTGTGTGTGTGTGTCTTCTTCCCTCTTTTGATTTTTGGCGCTGGAATAT	1240
Db	1757	-----CTTTTCATTCGCTTTGTTTGTTCGCCAAGACGAT	1725
Qy	1241	TTATTTATCATATTTCTTGAATGGGTAAACATCTTTAGATTGAAGTTTTTCTCCTAGC	1300
Db	1724	TAGTTTCTGTCTCTCTAGGGTATAGCTTGCCTCCTTATGTGGGCTTTACCAATTAT	1665
Qy	1301	CTTCTT--TAGGTCTGCAATTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAAT	1358
Db	1664	ATCCTTTGTATGTCTGATTTGTAGAAGATATGTCTAAATTTGGTTTTGTCTATGGAAT	1605
Qy	1359	GTCTTTCTTTCTCCAACTATTTGTGACAGAAGTTTTTCTAAGTCGATGTCTGGCCTGA	1418
Db	1604	ATCTTGGTTTCTCCATCTATGTTAAATTTGAGAGTTTTGCAGATACAGTAACCTGGGCTGG	1545
Qy	1419	CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGGCGCTTCTTACATTTTGAAGTT	1478
Db	1544	CATTTGTGTTCTCTTAGGGTCTGTATGACATCTGTCCAGGATCTTCTGGCCTTCATAGTT	1485
Qy	1479	TCTATTGGAAAAGTCAGCTGTAATCTTAATACATCTGCCCTTATATGTTTAATTTGGTCTTT	1538
Db	1484	TCT--GGCGAGAAGTCTGGTGTGATCTGATAGTCTGCCCTTATATGTACTT--GACCTT	1427
Qy	1539	TTTTCCCTTGCACTCTTTTAAATCTTCTTGTCTATCTTTTAGTGAATTCGATTATTA	1598
Db	1426	TTTTCCCTTACTGCTTTTAAATTTCTTCTTATTTTGTGGTTTTGGTGTTTGACATTA	1367
Qy	1599	TGCACTGTGGGAGTTTCTTTTCCGGTCCAAATCTATTGTGTGTTTTGTATGCTTCTTGTA	1658
Db	1366	TGTGACGGGAGTGTTCCTTTCTGTGTCCAACTCTATTTGGAGTTCTGTAGGCTTCTTGTA	1307
Qy	1659	CCTTGATPAGCATCTCTTTCTCAAGGTTAGAAATTTTCTTTTGGTTTTCTTGAAAA	1718
Db	1306	TGCGTATGGGTATCTCTTTTTTTTAGTTAGGTAAGGAGTTTTCTCTATGATTTTGTTTGAAGA	1247
Qy	1719	TATTTTCCCTTGTGTTTGAACCTG-----CCTTCTTCCCTTCCCTCTATTTCTT	1764
Db	1246	CATTTACTGTCTCTTTGAGCTGGGAGTCTTCACTCTCTCTATACCTATTATCTCTAGGT	1187
Qy	1765	TTGGTTTTTGCATAGTGTCTCTGGCTCTCGAATGTTTTATGCTGTGAATTTATTTAGACT	1824
Db	1186	TTGATCTTCTCATGTAGTCTCGAATTTCTGTATGTTTGGACAGTAGCTTTTTCGCT	1127
Qy	1825	TAACTATTTCTTTGACCAAGGATTCATTTCTTCTATCTTGTCTTCACTGCTGAGATTC	1884

Db 1126 TTACATTAATCTTTGACAGTTGAGTCAATGATTTCTATGGAAATCTCTGCTCTTGAGATTC 1067
Qy 1885 TCTCTTCTATCTCTGTATTTCTGTCAGTGGCTGTCTCTGAGGTTC -TGTGGGTTTC 1943
Db 1066 TCTCTTCAATCTCTGTATTTCTGTGTGATGCTGTGATCTAGGGCTCTCTCTCTCC 1007
Qy 1944 TTAATTTTTCATTTCCAGATTTCTTCACTTTGGGTTTGTATTAATCTCTATTCCTCA 2003
Db 1006 TTGGTTTCTATATCCAGGTTGTTTCCATGTTCTCTTCTGATGCTCTATTTTCCA 947
Qy 2004 CTTTCAGTCTCTGAATGTTTACTCATTTTCTCCAGATTTTACATTTTTCATAGGTTT 2063
Db 946 TTTTAAATCTCTCACTGTT-----TGAITGTGTTTCTCTGGAAT 905
Qy 2064 CTTTAAATGGATTTATCATTTCTCTTCAAGGACCTTTTATGAATTCATTAATATGATGT 2123
Db 904 CTTTCAGGGAATTTTGGCAATTCCTCTCTGTAGGCTTCTACTGTTCTCTAAGGGAGTTCT 845
Qy 2124 TAAGGTCTCT 2133
Db 844 TCACGTCTTT 835

RESULT 12

US-10-191-803-720/c
; Sequence 720, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10191.803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: Patent In Ver..2.1
; SEQ ID NO 720
; LENGTH: 6335
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 X53581
US-10-191-803-720

Query Match 25.8%; Score 588.5; DB 1; Length 6335;
Best Local Similarity 62.7%; Pred. No. 5.8;
Matches 1173; Conservative 0; Mismatches 560; Indels 137; Gaps 18;
Qy 292 TCTATTTCTCATTTCTATCTTGGCTCATTTTAACTCAGTAGTGTGTTTGGTTTCC 351
Db 2595 TTTCTTTCTTATTTCTCTTGGACAGGTTATCATTTAGTAGGCAATTTTCAGTTTC 2536
Qy 352 ATAAGTTTGTAAAGTTTCTGTGTTTCTGTTGTTTGTGTTTATCTAGATTTAAAGCTGT 411
Db 2535 AAGTATATATGGCATTCT-----TTCCTTGATTGTTTATTGAAGCAGTTTTCAGGCGT 2482
Qy 412 GGTGTCAGATAGGACATAGATATTTTCAATTTGCTTTTATCTCTCGAGACTTGTCT 471
Db 2481 GGTGTCGATAGCATGATGGGATTTATTTCTTTCTGTACCTCTGTGAGGCCCTTT 2422

Qy 472 TGTCTTTGAAATATGATTAATCAATTTTGGAG--AGTTTTCATAGGTCCTGACAAAGGTAC 529
Db 2421 TTTGACCAATATATAGTCAATTTTGGAGAAAGTACCATGAGGAGCTGAGAGAGGTAT 2362
Qy 530 AGTCTTTGTGTTTGG-TGAAATAGTCTGTAATAAT-CTTAGGTCCACTTGGTTTATGA 587
Db 2361 ATCCCTTTGCTTTAGGATAGAATGTTCTATAAATATCCGTTAAGTCCCTTTGGCTCATGA 2302
Qy 588 CATCAGTTAGTCCAGCATTTCTCTGTTTCGTTTTCGTTTTCGATGACCTAACTGTTTGA 647
Db 2301 CTTCTCTTAGTCTGTCTCACTCTCTGTTT-AAATTTCTGTTTCCATGATCTGTCCAATGAT 2243
Qy 648 GAGAAATGGGTATTTCAAGTAGCCCACTATCTGTGTGAGGT-CAATAATGATTTTACG 706
Db 2242 GAGAGTGGGTGTTGAAATCTCCCACTATTTATGTGAGGTGCAATGTGTGTTTGGC 2183
Qy 707 TGTAGTGTGCTGTTTATGAACTTGGTGACATTTGTTTGGTGCATAGACATTAGA 766
Db 2182 TTTAGTAAGGTTTCTTTTACATATAGTAGGTGCCCTTGTATTTGGGGCATAGATATTAGG 2123
Qy 767 ATTGCAATGTCCTCTTGGTGA-TTTTCTTTGATGCCCTATGTAGTATTTCTTCCAAATCT 825
Db 2122 ATTGAGAGTTCACTCTTGGTGGATTTTCTTTGATGATATGAAGTGTCTTCTTATCT 2063
Qy 826 CATCTGTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATATAAATGACATGTATCGGCT 884
Db 2062 TTTTGTAGTACCTTTAGTTAAATAATGATTTTATTTGATATTTAGAAATGGCTACTCCAGCT 2003
Qy 885 TGCTTTCTTAGGGCCATTTGCTTAGAATA-TCCTTTTCCATCTTTTACTCTAAGGTGATGT 943
Db 2002 TGCCTTCTCTGACCAATTTGCTTTGAAATATGTTTTCAGCCCTTTCACCTCTCAGGTAATGT 1943
Qy 944 CTATCCATG---GTAGGTTGTCTTTTGGATGACAGCATAGGATGGAATCTGTTTTCAT 1000
Db 1942 CTGTCTTTGTCTCAGGTGTGTTTCTCTAGGACAGCAAGATGCAAGGTCTCTGTTGCGT 1883
Qy 1001 ATCCATCTGTTTACCAGTACTTTTCTTAGAGAAATTAAGATCATTTAGGTCATGATGT 1060
Db 1882 ATCCAGTTTGTAAATCTATGCTTTTATTTGGGGAGTTAAG-----GCCATTTATGT 1831
Qy 1061 TGAGAAATATCAATCAGCAGTGTTTTGTGGATTTCTTTGTTATCTTGCACCTTTGGAAGTGT 1120
Db 1830 TGAGAGATATTAAGGAATAGTATTTATGCTTCCGTTTATATTCATATTTGGATGTGAGG 1771
Qy 1121 GT 1180
Db 1770 TTAGTTTGTGTG----- 1758
Qy 1181 GTGTCTGTGTGTGTGTGTGTGTGTGTCTCTCCCTCTTTTGAATTTTGGCCTGGAATPAT 1240
Db 1757 -----CTTTCATTCGCTTTGTTTGTGCAAGACCAT 1725
Qy 1241 TTATTTATTCATATTTTCTTGAATGAGGTAAACATCTTTAGATTTGAAGTTTCTCCTAGC 1300
Db 1724 TAGTTTCTGTCTTCTTCTAGGATAGCTTGCCTCTTATGTGAGGCTTTTACCATTAT 1665
Qy 1301 CTTCTT--TAGGTCTGCATTTGAACATAGATATTTCTTACATCTCATTTTATCTTAGAAT 1358
Db 1664 ATCCCTTTGAGTGTGGATTTGTAGAAAGATATGTGTAAATTTGTTTGTTCATGGAAT 1605
Qy 1359 GTCTTTCTTCTCCAACTATTGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGCGCTGA 1418
Db 1604 ATCTTGGTTTCTCCATCTATGTTAATTCAGAGTTTTTGCAGGATACAGTAACCTGGGCTGG 1545
Qy 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCAGGGCTTCTTACATTTTGAGTT 1478
Db 1544 CATTTGTGTCTCTTAGGCTCTGTATGACATCTGTCCAGGATCTTCTGGCTTCTATAGTT 1485
Qy 1479 TCTATTGGAAAAAGTCAGGTGTAAATCTAATACATCTGCTCTTTATATGTTAATTTGCTCTTT 1538
Db 1484 TCT-GGCGAAGATCTGGGTGTGATTTCTCATAGGTCTGCTTTATATGTTACTT-GACCTT 1427
Qy 1539 TTTCCCTGTCATCTTTTAAATATCTTCTTGTGTTCTA-ACITTTAGTGTATGTTGATTA 1598

Db 1426 TTTCCCTTACTGCTTTTAAATATCTTTCTTTATTTTGGGTTGGTGTGACTATTA 1367
Qy 1599 TGCATGTGGGAGTTCTTTCCGGTCCAAATCAATTTGGTGTGTTTGTATGCTTCTGTA 1658
Db 1366 TGTACGGGAGGTGTTCTTTCTGGTCCCAATCAATTTGGAGTTCTGTAGGCTTCTGTA 1307
Qy 1659 CTTGATAGGATCTTTCTCAAGTTTGAAGAAATTTTCTTTTGGTGTGTTTGTGAAA 1718
Db 1306 TGCCTATGGGTATCTTTTATAGTTAGGAAAGTTTCTTCTATGATTTTGTGAAGA 1247
Qy 1719 TATTTCCCTGCTTTTGACCTG-----CCTTCTCCCTTCTCTATTCCT 1764
Db 1246 CATTTACTGCTTTTGAGCTGGAGCTTTCACCTCTCTTATACCTATTAATCCTTAGGT 1187
Qy 1765 TTGTTTTTGCATAGTGTCTTGGCTTCTGGAATGTTTTATGCTTGGATTAATTTAGACT 1824
Db 1186 TTGATCTTCTCATTTAGGCTTGGATTTCTGTATGTTTGGACCAAGTGTCTTTCCGCT 1127
Qy 1825 TAAATTTTCTTACCAAGATGATCCATTTCTTCTATCTTGTCTCACTGCTGCTGAGATTC 1884
Db 1126 TTAATTAATCTTTGACAGTTGAGTCAATGATTTCTATGGAATCTTCTGCTTGGATTC 1067
Qy 1885 TCTCTTCTATCTTTGATTTCTGACGTGAGGCTTGTCTCTGAGGTTCC-TGTTGGGTTTC 1943
Db 1066 TCTCTTCCATCTCTGATTTCTGTTGGTGTGATCTGATCTAAGGCTCTTGTCTCTTCC 1007
Qy 1944 TTAATTTTCTTCAATTCAGATTTCTTCAAGTTTGGGTTTTGTTTAAATCTATTTCCA 2003
Db 1006 TTGGTTTTCTATATCCAGGTTGTTTCCATGTTCTTCTTCTGATTTCTTATTTCCA 947
Qy 2004 CTTTCAGGCTCTGAAATGTTTTTACTCATTTTCTCCAGATTTTACATTTTCTATGTTT 2063
Db 946 TTTTAAATCTTCAACGTGT-----TGATTTGTTTCTCTGGAAT 905
Qy 2064 CTTTAAATGATTAATTCATTTCTTCAAGGACCTTTTATGAATTCATAAATGATGT 2123
Db 904 CTTTCAGGATTTTGGGATTCCTCTGTAGGCTTCTACTTGTCTCTAAGGAGTTCT 845
Qy 2124 TAAGTCTCT 2133
Db 844 TCAGTCTTT 835

RESULT 13

US-10-374-077-207
; Sequence 207, Application US/10374077
; Publication No. US20040006779A1
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; Yu, Chang-En
; Oshima, Junko
; Mulligan, John T.
; Schellenberg, Gerald D.
; TITLE OF INVENTION: ANTIBODIES AGAINST GENE PRODUCTS RELATED TO
; WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/374,077
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Rosenman, Stephen
REGISTRATION NUMBER: 43,058
REFERENCE/DOCKET NUMBER: 100107.401D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 207:
SEQUENCE CHARACTERISTICS:
LENGTH: 29604 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 207:
US-10-374-077-207

Query Match 25.8%; Score 588.3; DB 1; Length 29604;
Best Local Similarity 62.9%; Pred. No. 3;
Matches 1170; Conservative 0; Mismatches 567; Indels 123; Gaps 18;
Qy 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTAGTGTGTTGTTTCC 351
Db 23722 TTTCTTTCTTTATTCCTTCCCTTGACCAAGGTATCATTTGAGAAGAGTGTATTTCAGTTTCC 23781
Qy 352 ATAAGTTTCTGAAATTTCTGTTTCTGTTTCTGTTTGTGTTTATCTAGATTTAAAGCTGT 411
Db 23782 ACGTGAATGTGGCTTTCCATTTATTTATGTTGTTATGAAGATC---AGCCTTAGGCCAT 23838
Qy 412 GGTGTCAGATAGGACATAGATTAATTTCAATTTGTTTCTTTATCTGTCGAGACTTGTCT 471
Db 23839 GGTGTCCTGATAGGATACATGGACAAATTTCAATATTTTGTATCTATTGAGGCCCTGTTT 23898
Qy 472 TGTTTTGAATATGATTTCAATTTTGGAGA---GTTTCATAGGCTGCTGACAGAGGTAC 529
Db 23899 TGTGACCAATTAATGTTCAATTTTGGAGAAGGCTCCGAGGTGCTGAGAAGAAGGTAT 23958
Qy 530 AGTC-TTTTGTGTTTGGTGAATAGTCTGTAATATCTCT-AGGTCACCTTGGTTTATCA 587
Db 23959 ATCCTTTTCTTTAGGATAAATGTTCTGTAGATATCTGTAGGTCATTTGTTTCATTA 24018
Qy 588 CATCAGTTAGCTCCAGCAATTTCTGTTTCTGTTTCTGTTTGTGAGATGACCTAACTGTTGA 647
Db 24019 CTTCTGTTAGTTTCACTGTGCTCCTGTTTAG-TTTCTGTTTCCAGCATCTGCTTTGAA 24077
Qy 648 GAGATGGGATTTGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGTCATTTTAC 706
Db 24078 GAAAGTGTGTGTGAGTCTCCACTATTTATGTTGAGGTGCAATGATGCTTTGAGC 24137
Qy 707 TGTAGCTGTGTTGTTTATGAACCTTGGGTGACATTTGTTTGGTGCATPAGACATTAAGA 766
Db 24138 TTTACTAAAGTGTCTCTAATGAATGTGGCTGCCCTTGCAATTTGGTGCCTAGATATTCA 24197
Qy 767 ATTGCAATGCTCCTTGGTGGATTTT-CCTTTGATGCCCTATGCTAGTATTCTTCCCAATCT 825
Db 24198 ATTGAGTGTCTCTTTGGAGGATTTTACCTTTGATGAGTGAAGTGTCCCTTCTTGTCT 24257
Qy 826 CATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATTAAATGACTGTATCGGCT 884
Db 24258 TTTTGTATAACTTTGGGTTTGAAGTCGATTTTATCCGATACTATAATGCTACTCCAGCT 24317
Qy 885 TGTCTTCTAGGCGCAATTTGCTTAGAATA-TCTTTTCCATCTCTTTTACTCTAAGGTGATGT 943
Db 24318 TGTTTCTTCAGTCCCAATTTGCTTGGAAATTTGTTTCCAGCCCTTTTACTCTGAGGTAGT 24377
Qy 944 CTATCCAT--GGTAGGTTGTCTTTTGGATGACAGCAGTAGGATGGATCTTGTGTTTCTAT 1000
Db 24378 CTGTCTTTTCCCTGAGATGGGTTTCTGTAGCAGCAGAAATGTTGGGCTCTGTTTGTGT 24437
Qy 1001 ATCCATCTGTTTACCAGATATCTTTTCTPAGAGAAATTAAGATCAATAGTCAATGATGT 1060
Db 24438 AGCCAGTCTGTTAGTCTATGTTCTTTTATTTGGAATTCAG-----TCCATTGATAT 24489
Qy 1061 TGAGAATTAATCAATGAGCAGTGTGTTGTGATTTCTTGTATCTTGCACITGTGAGAGTGT 1120

```
Db 24490 TAAGAGATATTAAAGAAAGTAATGTGCTCTCTTTTATTTTGTGTAGAGTTGGCA 24549
QY 1121 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1180
Db 24550 TTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 24582
QY 1181 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1240
Db 24593 -----TTGAATGAT 24591
QY 1241 TTATTAATCATTTCTTGAATGTGGTAACATCTTTAGATTGAAGTTTCTCTCTAGC 1300
Db 24592 TACTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 24651
QY 1301 CTTCCTTT--AGGTCTGCATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTAGAAT 1358
Db 24652 ATCCCTTGAAGGGCTGGATTCTGTGAAGATATTTGTGTGAATTTGTGTGTGTGTGT 24711
QY 1359 GTCCTTTCTCTCCAACTATTTGTGACAGAAAGTTTCTTAAGTGCAGTGTCTGGCTGA 1418
Db 24712 ACTTTGGTTCTCCATCTATGTAATTTAGAGATTTGGCTGTATAGTAGCTGGCTGG 24771
QY 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCTCTTACATTTTGAATT 1478
Db 24772 CATTTGTGTCTCTTAGTTCTGTATACATCTGTCCAGGGCTCTTCTGGCTTTCATAGTC 24831
QY 1479 TCTATTTGAAAAGTCAGGTGAATCTAAATACATCTGCCTTTATATGTTAAATGGCTTT 1538
Db 24832 TCTGCT-GAAAAGTCTGGTGAATTTCTGATAGGCTTCTTTATATGTTACTT-GACCTT 24889
QY 1539 TTTCCCTTGCATCTTTTAAATATCTCTTTTGTCTATATCTTTTGTAGTATGATTATTA 1598
Db 24890 TCTCCCTTACTGCTTTTAAATCTAATCTTTTATTTAGTGAATTTGTGTCTGATTTATTA 24949
QY 1599 TGCACTGTGGGAGTTCTTTTCCGGTCCAACTATTTTGTGTGTGTGTGTGTGTGTGTGT 1658
Db 24950 TGTGCGGGAGGAATTTCTTTCTCTGCTCAGTCTATTTGGAGTCTCTGTAGGCTCTTGTGA 25009
QY 1659 CTTGTATAGGATCTCTTCTCAAGGTTAGGAATTTTCTTTTGTGTGTGTGTGTGTGTGT 1718
Db 25010 TGATCAAGGATCTCTTTTCTTTTATGTTGGAAAGTTTCTCTATTTTGTGTGAAGA 25069
QY 1719 TATTTTCCCTGTTTGAACCTG-----CCTTCTTCCCTTCTCTATTTCTT 1764
Db 25070 TATTAGCTGGCCCTTAAAGTTGAAATCTTCAATCTCATCAATCTCTATTTATCGTAGT 25129
QY 1765 TTGGTTTTGATAGTGTCTCTGGCTTCTGGATGTTTATGCTGGAATTTATTTTAGACT 1824
Db 25130 TTGCTCTCTCTATTTGTGCTTCTGGAATTTACCTGGATTTTGTAGTTAGGATCTTTTGCAT 25189
QY 1825 TAAATTTTCTTTGACCAAGATATCCATTTCTTATCTTGTCTTCTACTGCTGAGATTC 1884
Db 25190 TTGATTTTCTTTGACTGTGTGTGCGATGTTCTATGGAATCTTCTGCACTGAGATTC 25249
QY 1885 TCTCTTATCTTGTATTTCTGTGATTTCTGTGAGGCTGTCTCTGAGGTTCTCTG-TTGGGTTTC 1943
Db 25250 TCTCTTCAATTTCTTGTATTTCTGTGTGTATGCTGCACTATGATGTTTCCAGATCTCTTTTC 25309
QY 1944 TTAATTTTCTCATTTTCCAGATTTCTCTCAGTTTGGTTTGTGTTTATTAATCTATTTTCCA 2003
Db 25310 CTAGATTTCTATCTCCAGGTTGCTGCTTGGTTTCTTTTGTGTGTGTGTGTGTGTGTGT 25369
QY 2004 CTTTCAGTCTCTGAAGTTTACTCATTTTCTTCCAGTAATTTTCTTCTTCTTCTTCTTCTT 2060
Db 25370 CTTTGTAGTCTAGTATGTTTGTGTCTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 25429
QY 2061 TTTCTTTTAAAGATTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2120
Db 25430 TTTCTTTTAAAGATTTCTACCTGTTTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 25489
```

```
US-10-322-281-115
; Sequence 115, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115
; LENGTH: 289190
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(289190)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-115
```

```
Query Match 25.8%; Score 587.8; DB 1; Length 289190;
Best Local Similarity 62.4%; Pred. No. 1.1;
Matches 1168; Conservative 0; Mismatches 582; Indels 122; Gaps 17;

QY 292 TCTATTCTTGTATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAGTTGTGTGTTC 351
Db 132550 TTTCTTTCTTTATTTCTTCTTCTTCCCTGACCAAGGTATCATTTGAGAAGTGTATTTCACTTTCC 132609
QY 352 ATAAAGTTGTAAAGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 411
Db 132610 ACCTGAATGTTGGCTTCTCCATTTATTTATGTTGTATTGAAGATC---AGTCTTAGGCCAT 132666
QY 412 GGTGCTCAGATAGGACATAGAGTATTTTCAATTTGTCTTTTATCTGTCGAGACTTGTCT 471
Db 132667 GGTGCTGATAGGATACATGGACAAATTTCAATATTTTGTATCTATGAGGCTGTCT 132726
QY 472 TGTTTTGAATATGTATTAATTTTGGAGA--GTTTCTAGGGTCTGACAGAGGTTAC 529
Db 132727 TGTGACCAATATATGTTCAATTTTGGAGAAGTCCCTGAGGTCTGAGAAGAGGTAT 132786
QY 530 AGTC--TTTGTGTTTGTGGAATAGTCTGTAAATATCT-AGTCCACTTGTGTATGA 587
Db 132787 ATCCCTTTGTTTGTAGATAAAATGTTCTGTAGATATCTGTGAGGTCCATTTGTTTCATAA 132846
QY 588 CATCAGTTAGTCTCCAGCATTTCTCTGTTTGTGTTTGTGAGATGACTTAACCTTTGGA 647
Db 132847 CTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 132905
QY 648 GAGATGGGGATTTGAAGTAGCCCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 706
Db 132906 GAAAGTGTGTGTGTGAAGTCTCCCACTATTTATTTGTGAGGTGCAATGTATGCTTTGAGC 132965
QY 707 TGTAGCTGTGCTGTTTGTATGAATTTGGTGACATTTGTGTTGTGTGTGTGTGTGTGTGTGT 766
Db 132966 TTTTCTAAAGTGTCTTAAATGAATGTGCTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 133025
QY 767 ATTGCAATGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 825
Db 133026 ATTGAGAGTCTCTTGGAGGATTTTACCTTTGATGAGTATGAAGTGTCTCTCTGTGTCT 133085
QY 826 CATCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 883
Db 133086 TTTTGTGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133145
QY 884 TTGCTTCTTGTAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 942
Db 133146 TTTGTTTCTTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 133205
QY 943 TCTATCCAT-----GGTAGGTGTCTTTTGTGATGACGAGTAGGATGATGATGATGATGAT 998
Db 133206 TCTGTCTTTTCTCCCTGAGATGGGTTTCTGTGTAAGACAGAAATGTTGGGTCTCTTTGT 133265
```

QY 999 ATATCCATTCGTATACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATTTGAT 1058
Db 133266 GTAGCCAGTCTGTAGTCTATGCTCTTTTATTTGGGGAATTGAG-----TCCATTTGAT 133317
QY 1059 GTTGAGAATTAATCAATGAGCAGTGTCTGTGATCTCTGATCTTGCACCTTGTGAAGTGT 1118
Db 133318 ATTAGAGATATTAAGGAAAGATTAATGTTGCTCCCTTTAATTTTGTGTAGAGTTGG 133377
QY 1119 GT 1178
Db 133378 CATTCGT 133412
QY 1179 CTGT 1238
Db 133413 -----TTGAATG 133419
QY 1239 ATTTATTAATCATATTTTCTTGAATGGTGAACATCTTTAGATTGAAGTTTTTCTCCTA 1298
Db 133420 ATTACTTTCTTGGTTGTCTAGGGCGTGAATTCGTCCTGTATGTCTCTTTTCTGTGA 133479
QY 1299 GCCTTCCTT--AGGTCTCATTTGAAGATAGATATCTTTACATCTCAATTTATCTTGA 1356
Db 133480 TTATCTTTGAAGGCTGTGATTCGTGGAAAGATATTTGTGAACCTTGGTTTTGTCTGGA 133539
QY 1357 ATGTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTTTCTAAGTCAGTAGTCTGCGCT 1416
Db 133540 ATACTTGGTTCTCCATCTATGTAAATTGAGATTTGGCCGGTATAGTAGCCCTGGCT 133599
QY 1417 GACATCTGTAGTCTCTTGGATCTGTAGCAATCTGTGACGGGCTCTTACATTTTGA 1476
Db 133600 GGCATTTGTGTCTCTTAGTCTGTATAACATCTGTCCAGGCTCTTCTGGCTTTCATAG 133659
QY 1477 TTTCTATTGGAAGTCAAGGTGAATCTATATACATCTGCTTTATATGTATTTTGTGTCT 1536
Db 133660 TCTCTGTT-GAAAGTCTGTGTATTTCTAGAGCCCTCTTTATATGTACTT-GACC 133717
QY 1537 TTTTCTCCCTGCATCTTTAATATCTTTCTTCTTATATCTTATCTTATCTTATCTTAT 1596
Db 133718 TTTCTCCCTTACTCTTTAATATCTTATCTTATCTTATCTTATCTTATCTTATCTTAT 133777
QY 1597 TATGCACTGTGGGAGTTCTTTTCCGGTCAATCTATTTGGGTGTGTGTGTGTGTGTGT 1656
Db 133778 TATGTGTGGGAGAAATTTCTTTTCTGTCCAGTCTATTTGGAGTTCTGTATGCTTCTG 133837
QY 1657 TACCTTGATAGGCATCTTTCTCAAGGTAGGAAATTTTCTTTTGTGTGTGTGTGTGTGAA 1716
Db 133838 TATGATCATGGGATCTCTTTTATATGTTTGGAGATTTTCTTCTATTAATTTTGTGAA 133897
QY 1717 AATATTTCCCTGCTTTGACCTG-----CCTTCTCCCTTCCCTCTATTC 1762
Db 133898 GATATTAGCTGGCCCTTTAAGTTGAAATCTTATCTCATCAATTCCTATTATCCGTAG 133957
QY 1763 CTTTGGTTTTTGCATAGTCTCTGGCTTCCGTGATGTTTATGCTCGGATTAATTTTGA 1822
Db 133958 GTTTGGTCTCTCATTTGTCTGTGATTTACCTGGATGTTTGTAGTTAGGATCCCTTTTGA 134017
QY 1823 CTTAACAATTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGAT 1882
Db 134018 TTTTGTATTTCTTTGACTGTGTGTGCAATGTTCTCTATGGAATCTTCTGACCTGAGAT 134077
QY 1883 TCTCTCTCTATCTCTGTATCTGTGAGGAGCTGTCTGAGGTTCTG-TTGGGT 1941
Db 134078 TCTCTCTCCATTTCTTGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 134137
QY 1942 TCTTAATTTTCTATTTCCAGATTTCTCTAGTTGGTTTTGTGTGTGTGTGTGTGTGTGTGT 2001
Db 134138 TCCTAGGTTTCTATCTCCAGGTTGCTCGTTTGGTTTTCTTTATGTGTCTACTTC 134197
QY 2002 CACTTCCAGTCTGAAATGTTTACTCATTTTCTCCAGATATTTTACATTTTCAATAGGT 2061
Db 134198 CCTTTTGTAGTCTAGTATGTTTGTGTCAATTCATCCAGTGTGTGTGTGTGTGTGTGTGTGT 134257

QY 2062 TTCTTTAATGGATTTATTCATTTCTCTCAAGACCTTTTATGAATTCATATAAATGTAT 2121
Db 134258 CTTTCTTTAAGACCGCTGTAACCTTTTAGCAGTCTCTCTCTGTAATCTTTTAGTGACTT 134317
QY 2122 GTTAAGGTCCTT 2133
Db 134318 ATGAAAGTCCTT 134329

RESULT 15

US-10-175-523-57/c
; Sequence 57, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/10795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57
; LENGTH: 225883
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-57

Query Match 25.7%; Score 585.399; DB 1; Length 225883;
Best Local Similarity 63.6%; Pred. No. 1.2; Indels 120; Gaps 17;
Matches 1131; Conservative 0; Mismatches 526;

QY 292 TCTATTTCTTGATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGTTTCC 351
Db 187754 TTTCTTTCTTTATCTCTTCCCTTGACCAAGGTATCATTGAGAAGAGTGTGTTTCAGTTTCC 187695
QY 352 ATAAAGTTTCTAAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 187694 ATGTGAATGTTGCTTTTCCATTTATTTATGTTGTTTATTGAAGAT---CAGTCTTAGGCCAT 187638
QY 412 GGTGGTCAGATAGGACATAGAGTATTTATTTCAATTTGTTCTTTTATCTGTGAGACTTGCTT 471
Db 187637 GGTGGTCTGATAGGATACATGGACATTTCAATTTTGTATCTGTGAGCCCTGTTT 187578
QY 472 TGTTTTGAATATGTAATTCATTTTGGAGA--GTTTCATAGGGTGTCTGACAGAAGGTAC 529
Db 187577 TGTGACCAATTAATGTTCAATTTTGGAGAAGTCCCATGAGTGTCTGAGAAGAAGGTAT 187518
QY 530 AGTC-TTTCGTGTTTGGTGAATAGTCTGTAATAATCTCT-AGGTCCTACTTGGTTATGA 587
Db 187517 ATCCTTTGTGTTTAGGATAAATGTTCTGTAGATATCTGTAGGTCCTATTTGTTTCAATA 187458
QY 588 CATCAAGTAGCTCCAGCAATTTCTCTGTTTCTGTTTGTGTTGAGATGACCTAACTGTTGA 647
Db 187457 CTTCTGTTAGTTCATCTGTGTCCCTGTTTAG-TTTCGTGTTTCCACGATCTGTCCATTGAT 187399

QY 648 GAGAAATGGGATTTGAAGTAGCCCACTATCTGTGTGAGGT-CAATATATGTGATTTAGC 706
Db 187398 GAAAGTGGTGTGTTGAAGTCTCCCACTATTATTTGTGTGAGGTGCAATGTGTGTTGAGC 187339
QY 707 TGTAGCTGTGCTGTTTATGAACCTTGGGTGACATCTGTTGGTGCATAGACATTAAGA 766
Db 187338 TTTACTAAAGTGTCTTAAATGAATGGCTGCCCTTGCAATTTGAGGCGTAGAATTCAGA 187279
QY 767 ATTGCAATGTCTCTGTGGTGATTTT-CCTTTGATGCTATGTAGTATTTCTTCCCAATCT 825
Db 187278 ATTGAGAGTTCATCTGGAGGATTTTACCCTTTGATGAGTAGAATGAAGTGTCCCTCTCTGTCT 187219
QY 826 CATCTGCTAGTTTGGTGTAAAGTCTA-TTAGTTCAGATATTAAGATGACTGTATCGGCT 884
Db 187218 TTTTGTATAACTTTGGGTGGAGTCCGATTTTATCCGATATTAAATGGGTACTCCAGCT 187159
QY 885 TGCTTCTTAGGGCCATTTGCTTAGAATA-TCTTTTCCATCCTTTTACTCTAAGTGTATGT 943
Db 187158 TGTCTTTCAGATCATTTGCTTGGAAATTTGTTTTCAGCCTTTTATCTCGAGGTAGTGT 187099
QY 944 CTATCCAT- --GGTAGGTGTCTTTTGGANGCAGCAGTAGGATGGAATCTGTGTTTCAT 1000
Db 187098 CTGTCTTTTCCCTGAGATGGTTCCTGTAAAGCAGCAGAAATGTTGTCTCTGTCTGTGT 187039
QY 1001 ATCCATCTGTATCCAGTATCTTTTCTTAGAGAAATTAAGATCATTTGAGTCAATGATGT 1060
Db 187038 AGCCAGTCTGTATGTCTATGTCTTTTATTTGGGAATGAG- -----TCCATGTAT 186987
QY 1061 TGAGAAATPATCAATGAGCAGTGTGTTGGATCTTGTATCTTGCACTGTGGAAGTGTGT 1120
Db 186986 TAAGAGATTAATAGAGAAAGTAATTTGTGCATCTTTTATTTTGTGTTACAGTTGCCA 186927
QY 1121 GT 1180
Db 186926 TTCTGTCTTGT 186894
QY 1181 GT 1240
Db 186893 -----TTGAGTGTAT 186885
QY 1241 TTATTAATCAATTTCTTGAATGTGGGTAAATCTTTTAGATTTGAAATTTTCTCTAGC 1300
Db 186884 TACTTCTGT 186825
QY 1301 CTCTCTTT- --AGTCTGCAATTTGAAGATAGATATTTTACATCTGATTTATCTTACAA 1358
Db 186824 ATCCCTTTCATGGCTGGATTTCTTGAAGATPATCGTGTGAATTTGGTTTGTCTGTGAAT 186765
QY 1359 GTCTTTCTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCTGA 1418
Db 186764 ACTTTGGTTTCTACATCTATGTAATTTGAGAGTTTGGCCGGTATAGTAGCCTGGGCTGG 186705
QY 1419 CATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGAGGCTCTTACATTTTGGATTT 1478
Db 186704 CATTTGTGTCTCTTGTAGTCTGTATACATCTGTGAGGCTCTCTGCGCTTTCATAGTC 186645
QY 1479 TCTATTGGAAGAGTCAGGTGTAATTTCTAATACATCTGCTTTATATGTTAATTTGCTTT 1538
Db 186644 TCTGTG- GAAAGTCTGTGTGAATTTCTGATAGGCTTCTTTATATGTTACTT- GACCTT 186587
QY 1539 TTTCCCTGTGCATCTTTTAAATTTCTTTCTTTGTTCTATACATTTTAGTGAATTTGATTA 1598
Db 186586 TCTCCCTTACTGCTTTTAAATTTCTATCTTTTATTAGTGCATTTGTTGTTCTGATTAATTA 186527
QY 1599 TGCATGTGGGAGTCTTTTCCGGTCCCAATCTATTGTTGTTGTTTGTATGCTTCTTGTGA 1658
Db 186526 TGTGTGGGAGTGAATTTCTTTTCTGGTCCAGTCTATTGAGATCTGTAGGCTTCTGTA 186467
QY 1659 CTTGATAGGATCTCTTTCTCAAGGTTAGAAATTTTCTTTTCTTTTCTTTTCTTTGTTT 1718
Db 186466 TGATCATGGCATCTCTTTCTTTATGTTTGGAAAGTTTCTTCTATTATTATTGTTGAAGA 186407
QY 1719 TATTTTCCCTTGTGTTTGTACCTG- -----CCTTCTTCCCTCTCTCTATTCTT 1764

Db 186406 TATTAGCTGGCCCTTTAAAGTTGAAAAATCTTCATTTCAATCAATTCCTATTATCTGTAGGT 186347
QY 1765 TTGGTTTGTGATAGTGTCTCTGGCTTCTCTGGAATTTTTATGCCCTGGATTTATTTAGACT 1824
Db 186346 TTGGTCTCTCTCANTGTGTCTGGATTTACCTGGATGTTTGGATGATCTTTTGCATT 186287
QY 1825 TAAATTTTCTTTTGAACCAAGGATTCATTTCTTCTATCTGTCTTCACTGCTGAGATTC 1884
Db 186286 TTGTATTTTCTTTGACTGTGTGTGCAATGTTCTCTATGAAATCTTTCTGCACCTGAGATTC 186227
QY 1885 TCTCTTCTATCTCTTGTATTTCTCTCAGTGAGGCTGTCTCTGAGGTTCTCTG-TTGGGTTTC 1943
Db 186226 TCTCTTCCATTTCTTGTATTTCTGTTGCTGATGCTGCAATCTATGTTCCAGATCTCTTC 186167
QY 1944 TTAATTTTCTTATTTCCAGATTTCTTCTCAGTTTGGGTTTGTATTATTAATTTCTATTCCA 2003
Db 186166 CTAGGTTTCTATCTCTCCAGGTTGCTTGTCTTGGGTTTCTTTATTTGTCTACTTCCC 186107
QY 2004 CTTTCAAGTCTGAAATGTTTACTCATTTTCTCTCCC 2040
Db 186106 CTTTGTAGTCTAGTATGTTTGTTCATTTCCATCAC 186070

RESULT 16
US-10-003-806-10/c
; Sequence 10, Application US/10003806
; Publication No. US20020119229A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulnik, Alexander I.
; APPLICANT: Zhu, Qichao
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066US1/10024824
; CURRENT APPLICATION NUMBER: US/10/003,806
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,872
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 173808
; TYPE: DNA
; ORGANISM: Mouse
US-10-003-806-10

Query Match 25.7%; Score 585.1; DB 1; Length 173808;
Best Local Similarity 62.8%; Pred No. 1.4;
Matches 1188; Conservative 0; Mismatches 569; Indels 123; Gaps 18;
QY 292 TCTATTCTTGATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGTTTCC 351
Db 123632 TTCTTTCTTTATTCCTTCTGACCAAGGTATCATTTGAGAAGAGTGTGTTTCACTTCC 123573
QY 352 ATAGTTGTGTAATTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 123572 ATGTGAATGTTGGCTTCTGTTTATTTATTTGTTGTTGTTGTTGTTGTTGTTGTT 123516
QY 412 GGTGTCAGATAGGACATAGAGTATTTTCAATTTGTTTATCTGTCGAGACTTGTCT 471
Db 123515 GGTGTCAGATAGGACATAGGACAAATTTCAATTTTGAATCTGTGAGGCTGAT 123456
QY 472 TGTGTTGAAATATGTAATTTCAATTTTGGAGA- -GTTTATAGGCTCTGACAGAGTAC 529
Db 123455 TGTGACCTATTATGTTGGTCAATTTTGGAGAAGGTACCAGGTGCTGAGAGAAGGTAT 123396
QY 530 AGTC-TTTCGTTTGTGGTGAAATAGTCTGTAATATCTCT-AGGTCCACTTGGTTTATGA 587
Db 123395 ATCCCTTTGTTTGTGGTGAATATGTTCTGTAGATATCTGTGAGATCTTGTGTTTCA 123336
QY 588 CATCAGTTAGTCCAGCATTTCTCTGTTTGTGTTTGTGTTGAGATGACCTAACTGTTGA 647
Db 123335 CTTCTGTTAGTTCAAGTGTGTCCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 123277

Qy 1719 TATTTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCTCTATTCTT 1764
Db 122284 TGTTCCTGGACCTTTTGAGTTGAAATCTTCACTTCACTCCACTTATTATTCOGTACCT 122225
Qy 1765 TTGGTTTTGCAATAGTGTCTCTGGCTTCTGGATGTTTTATGCTCGATTTATTAGACT 1824
Db 122224 TTGGCTTTCTTATTGTTGCTGGAATTCCTGATATTTTGGATAGGATCTTTTGCATT 122165
Qy 1825 TAACTTTTCTTTGACCAAGGATCAATTTCTTCTATCTTTGCTTCACTGCTCGAGATTC 1884
Db 122164 TTCCATTTTCTTTGATTTGTCGCAATGTTCTCTATGGAATCTTCTGACCTGAGATTC 122105
Qy 1885 TCTCTTCTATCTTCTTATTTCTGTCAGTCACTGAGCTTCTCTGAGTTCTCTG-TTGGGTTT 1943
Db 122104 TCTCTTCCATCTCTTATTTCTGTTCTGCTGATGCTCAATCTATGTTTCAGATTTCTTTTC 122045
Qy 1944 TTAATTTTCTTCAATTTCCAGATTTCTTCAATTTGGGTTTTGTTTATTAATTTCAATTTCCA 2003
Db 122044 CTAGGCTTTCTATCTCTAGTGTGCTGCTTGGATTTTCTTTTATTTGCTCTACTTCC 121985
Qy 2004 CTTTCAGTCTGGAATGTTTTACTCAATTTTCTCTCC-----AGTATTTACATTTTCATAGG 2060
Db 121984 TTTTATAGTCTAGTATGTTTGTTCATTTCCATTCACCTGTTTGTATGTTTTTCTCTCTT 121925
Qy 2061 TTTCTTTTAAATGGAATTTATTTCAATTTCTCTTCAAGGACCTTTTATGAATTCATAAAATGTA 2120
Db 121924 TTTCTGTAAGACTTCTACTCTGTTGATGTTGTTTTCTGTTTTCTTTTAAAGGACTTGT 121865

RESULT 17
US-10-322-281-271
; Sequence 271, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 271
; LENGTH: 120239
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(120239)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-271

Query Match 25.5%; Score 580,301; DB 1; Length 120239;
Best Local Similarity 61.9%; Pred. No. 1.7;
Matches 1199; Conservative 0; Mismatches 602; Indels 135; Gaps 19;
Qy 292 TCTATTTCTTGAATTTCTATCTTGGCTCATTTTAACTCAGTAGTAGTGTGTTGGTTTCC 351
Db 21580 TTTCTTTCTTTATTCCTTCTTGGACGAGGTATCATTTGAGAAGAGTGTGTTTCAGTTTCC 21639
Qy 352 ATAGTTTCTAGTTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTAGATTTAAGCTGT 411
Db 21640 ACCTAAATGTTGACTTCTTATTTATTTATTTGTTTATTTGAAGATC---AGCCTTAGTCCAT 21696
Qy 412 GTGGTTCAGATAGGACATAGATTTATTTCAATTTCTTTTATCTGTCGAGACTTGTCTT 471
Db 21697 GGATGATCTGATAGGATGATGGAATTTCAATATTTTGTATCTGTTAAAGCCCTGTTT 21756
Qy 472 TGTTTGAAATATGATTTCAATTTTGGAGA--GTTTCATAGGTCGTGACAGAAGGTAC 529
Db 21757 TGTGACCAAGTATGATCAATTTTGGAGAAGTAGCATGAGGTATGAGAAGAGGTAT 21816
Qy 530 AGTCTTTGTTGTTTTGTTGGAATAGTCTGTAAATATCT-CTAGTCCACTTGGTTTATGAC 588

Qy 648 GAGAAATGGGTATTGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGATGATTTAGC 706
Db 123276 GAAAGTGTGTGTTGAAGTCTCCCACTATTATTGTGTGAGGTGCAATGTGTGCTTTGAGC 123217
Qy 707 TGTAGCTGTGCTTGTATTGAATCTGGGTGACATTTGTTTGTGTCATAGACATTAAGA 766
Db 123216 TTTACTAAAGTTCTTTAGTGAATGTGGCTCTCTGTTGATTTGGAGCATAGATTCAGA 123157
Qy 767 ATTGCAATGTCCTCTTGTGATTTT-CCTTTGAAGCTATGTTAGTATTTCTTCCCAATCT 825
Db 123156 ATTGAGAGTTCTCTTGGAGGATTTTACCTTTGATGAGATGAAGTGTCCCTCTTGTCT 123097
Qy 826 CATCTGCTTAGTTTTGGGTTTTAAGTCTAT-TAGTCAGATATTAATAGCTGTATCGCT 884
Db 123096 TTTTGTATGACTTTGGGTGGAAGTCAATCTTATCAGATATTAGATGGCTACTCCCTGCT 123037
Qy 885 TGCCTTTTAGGGCCATTTGCTTTAGAATA-TCTTTTCCATCTCTTTTACTCTAAGGTGATGT 943
Db 123036 TGTTCCTACACCATTTGCTTTGGAATTTGTTTCCAGCTTTTCAATCTCAGGTAGTGT 122977
Qy 944 CTATCCAT---GGTAGTTGCTTTTGGATGAGAGTAGGATGATCTGTTTTCAT 1000
Db 122976 CTATCTTTTCTCTGAGATGATTTCTCTGTAGCAGCAAAATGTTGGGCTCTTGTGTTGT 122917
Qy 1001 ATCCATTTCTGTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTCAGTCAATTGATGT 1060
Db 122916 AGCCAGTTTGTAGTCTATGCTTTTATTTGCGAGTTGAGA-----CCATGATGT 122865
Qy 1061 TGAGAAATATCAATGAGCAGTGTGTGGAATCTTTGTTATCTTGCATCTGTGAAAGTGTGT 1120
Db 122864 TAAAGATATTAAGAAAAAGTAATGTTGCTCTCTGTTATTTTATTTAGTTTAAAGGTGCA 122805
Qy 1121 GTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTCT 1180
Db 122804 TTTGTTCTTTGGGCTGCTCTCTTTAGTTTG-----122772
Qy 1181 GTGTGCTGTGCTGTGCTGTGCTGTCTCTCCCTCTTTTGAATTTTGGCTGGAATAT 1240
Db 122771 -----TTGAGGGAT 122763
Qy 1241 TTATTTATCAATTTCTTGAATGTTGGTAACATCTTTAGATTGAAGTTTTTCTCTAGC 1300
Db 122762 TACCTTCTGTTTTCTTAGGCGTTGTTCCGCTCTGTTATGTTGTTTTTCTGTTATT 122703
Qy 1301 CTCTCTTT--AGTCTGCAATTTGAGATAGATTAATCTTTTACATCTGATTTATCTTAGAAT 1358
Db 122702 ATCTTTGAAAGGCTGGATTCGTGGAGATTAATGCGTGAATTTGGTTTGTCTGGAT 122643
Qy 1359 GTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTTTCTAAGTCAGTAGTCTCGCCTGA 1418
Db 122642 ACTTTGGTTTTCTCCCTCTATGATAATGAGAGTTTGGCTGGGTATAGTAGCCTGGCTGC 122583
Qy 1419 CATCTGTAGTCTTTGGAGTCTGTAGACATCTGTGAGGCTCTTACATTTTGAATT 1478
Db 122582 AGTTTGTGTTCTTTAGTGTCTGTATAACATCTGTCCAGGCTCTCTGCTTTTCATAGTC 122523
Qy 1479 TCTATTGGAAGAGTCAGGTGTAATTTCTAATACATCTGCTTTTATATGTTAATTTGGTCTTT 1538
Db 122522 TCTGCT-GAAAACTGCTGTAATCTCATAGAGCTTGCCTTTATATGTTACTT-GACCTT 122465
Qy 1539 TTTCCCTGCACTTTTAAATCTTTCTTTTCTTTGTTCTATACCTTTTAGTGATTTGATTA 1598
Db 122464 TTTCCCTTACTGCTTTTAGTATCTTATCTTTTATTAGTGCAATTTGATGTTCTGATTA 122405
Qy 1599 TGCAGTGGGGAGTTTCTTTTCCGTTCCAACTATTTTGGTGTGTTTGTATGCTTCTTTGTA 1658
Db 122404 TGTGTCGGAGGAATTTCTTTTCTGCTCCAGTCTATTTGGAGTCTGTGAGCTTCTTGTGA 122345
Qy 1659 CTTTGATAGGCATCTCTTTCTCAAGGTTAGAAATTTTCTTTTGTGTTTCTTTGAAAA 1718
Db 122344 TGTTCATATGATCTCTCAATCTTTAGATTGGGAAGTTTCTTCAATAATTTTGTGGA 122285

Db 21817 ATCCCTTTGTTTGGGATAAAATGTTCTGTAGATATCTGTTACATCCATTTTTCATAAAC 21876
QY 589 ATCAGTTAGTCCACATTTCTCTGTTTGGTTTGGTTGAGATGACCTACTGTTGGAG 648
Db 21877 TTATGTTAGTTTCACTATGTCCTGTTTAG-TTCTATTTCCATGATTTGTCATTTGATG 21935
QY 649 AGAATGGGATTTGAAGTAGCCCACTATCTGTGTAGGT-CAATATGATTTTATGCT 707
Db 21936 AAAGTGGTGTGTTGAATCCCCCACTATTTATGTTGAGGTGCAATGTGCTTTTGGCT 21995
QY 708 GTAGCTGTGCTGTTTATGACTGGGTGACATGTTGTTGGTGATAGACATTAAGAA 767
Db 21996 TTACTAAAGTGTCTTTAATGAATGTTGCTGCTGTTGATTTGGAGCATAGATATTAGAA 22055
QY 768 TTGCAATGTCTCTTGGTGGATTTTCTTTGATGCTATGTAGTATTTCTCCCAATCTCA 827
Db 22056 TTGAGAGTTCTCTTTGGAGGATTTTATCTT-TGAGTATGAAGTCCCTCTTTGTCTTT 22113
QY 828 TCTGCTTAGTTTGGGTTTAAAGTCTATTAG-TCAGATATTAATGACTGTATCGGCTTG 886
Db 22114 TTGATGACTTTGGGTTGGAAGTCAATTTTATCGATATTAAGATGGCTACTCCAGCTTG 22173
QY 887 CTTCTTAGGGCAATTTGTTAGATA-TCTTTTCCATCTTTTACTCTTAAGGTGATGCT 945
Db 22174 TTTCTTCAGCAATTTGTTTGCATAATTTGTTTCCAGCTTTCTACTCTGAGTAGTGTCT 22233
QY 946 ATCCAT---GGTAGTGTCTTTTGGATGACAGTAGATGGAATCTTGTGTTTTCATAT 1002
Db 22234 GTCCTTATCCCTGAGATGGGTTTCCGTGAAGCAGCAGAACTGTTGTCTCTTTGTGTAG 22293
QY 1003 CCAATCTGTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGACTCATGATGTTG 1062
Db 22294 CAGTTTGTATGCTATGCTCTTTTATTTAGGGAGTTGAG-----TCCATTTGATGTTA 22345
QY 1063 AGAATATCAATGACAGCTGTTTGGATCTCTGTTATCTGACCTTTGTAAGTGTGTGT 1122
Db 22346 AGAGATATTAAAGAAAGTAATTTGTTGCTTCCATTAATTTT---TTTGTAAAGTTGCA 22402
QY 1123 GTGCTGT 1182
Db 22403 TTTCTGTTCTGTGGCTGTCTTTTATAGGTTG----- 22435
QY 1183 GTCGTGTCTGTGTGTGTGTCTCTCCCTCTTTTGTATTTTGGCCTGGAATTAATTT 1242
Db 22436 -----TTGAGAGATTA 22446
QY 1243 ATATATCATATTTCTGAAATGGGTAACATCTTTAGATGAAGTTTCTCTCCAGCT 1302
Db 22447 CTTTCTTACATTTCTAGAGTGTGTTTCTGCTCTGTTATTTGTTTCTGTTATAT 22506
QY 1303 TCTTTT-AGGCTGATTTGAAGTAGATATTTTACATCTGATTTTATCTTTAGAAATG 1360
Db 22507 CTTTGAAGGACTGGAATTCGTGGAAGATAATGTTGTAATTTGTTTGTGCTGGAATAC 22566
QY 1361 CTTTCTTTTCTCAACTATTTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCTGACA 1420
Db 22567 TTTGGTTTCTCCATCTATGTAATTTGAAAGTTTGGCTGGGTATAGTAGCCTGGCTGGCA 22626
QY 1421 TCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGAGGGCCTCTTACATTTGAGTTTC 1480
Db 22627 TTTGTGTTCTCTAGTGTCTGTATAACATCTGTCCAGATCTTCTGGCTTTTCATAGTCTC 22686
QY 1481 TATTGGAAAGTCAAGTGTAACTTCAATACATCTGCCCTTTATATCTTAATTTGCTTTTT 1540
Db 22687 TGGT-GAAAGTCTTGTGTAATCTGTATAGGCTGCTTTATATGTTACTT-GACCTTTT 22744
QY 1541 TCCCTTGCATCTTTTAAATTTCTTTTGTGTCTATCTTCTTGTAGTGAATTTATATG 1600
Db 22745 CCCCTTACTGTTGTAATTTATATCTTTTATTTAGTGCATTTGTTGTTCTGATTTATATG 22804
QY 1601 CACTGTGGGAGTTTCTTTTCCGGTCCCAATCTATTGTTGTTTGTATGTTCTTTGTACC 1660

Db 22805 TGTGGGAGAAATTTCTCTTCTGGTTCACTATTTGGAGTTCTGTAGGCTTCTGTATG 22864
QY 1661 TTGATAGGCATCTCTTTCTCAAGGTTAGGAAATTTTCTTTTGGTTTCTTGAATAA 1720
Db 22865 TTATGGGCATCTCTCTCTTTAGGTTTGGAAAGTTTCTTCTATATAATTTTGTGACGATA 22924
QY 1721 TTTTCCCTGCTTTTGACCTG-----CCTTCTCCCTTCTCTATTTCCCTTT 1766
Db 22925 TTTGCTGGCCCTTTAAGTTGAAATCTTCAATCTCATCTACTCCCATATATCCGTAGGTT 22984
QY 1767 GGTTTTGTATAGTGTCTCTGCTTCTCGATGTTTATGCTCGGATTTATTTAGACTTA 1826
Db 22985 GGTCTTCTGTGTGCTCTGGATTTCTCGATGTTTAAAGTTAGGATCTTTTGGCATTT 23044
QY 1827 ACATTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATCTC 1886
Db 23045 GCATTTCTTTGATTTGTTGCTGCGATGTTCTCTATGGAATCTTCTGTACCTGAGATCTC 23104
QY 1887 TCTTCTATCTCTGTATTTCTGTCAGTGGCTGTCTCTGAGGTTCCCTG-TTGGGTTCTT 1945
Db 23105 TCTTCCATTTCTGTATTTCTGTGTGCTGATCTCACTATGTTCCAGATTTCTTTCT 23164
QY 1946 AATTTTTCATTTCCAGATTTCTCTTCAAGTTTGGGTTTGTATTATTAATCTATTTCCACT 2005
Db 23165 AGGTTTCTATCTCCAGGTTGCTTCCCTCACCTTTGGGTTTCTTATTTGTGTCACCTCTCT 23224
QY 2006 TTCAGTCTCTGAAATGTTTACTCAATTTTCTCTCCAGATTTTACA-----TT 2052
Db 23225 TTAGTCTAGTATGTTTGTTCATTTTCCATCACCTGTTGGGATGTTTTCCTGTTT 23284
QY 2053 TTCATAGGTTTCTTAAATGGAATTTATCTTCTCTTCAAGGACCTTTTATGAATTCAT 2112
Db 23285 TTTCTTAAAGACTTGTAACTCTTTAGCAGTGTCTCTCTGATTTCTTTAAGTGAGTTAT 23344
QY 2113 AAAATATGTTAAGTCTCT-TGCCTTGTGCTTCAAGTATGTTGCAATCTCTCAGGCTAT 2171
Db 23345 AAAGTCTTCTTGTATGCTCTCTACCATCATCATGATATGCCCTTTAAATCCAGTCTAG 23404
QY 2172 TGTAAATAGGTTTATAG 2187
Db 23405 CTTTGGGTTTGTG 23420

RESULT 18
US-10-003-806-10
; Sequence 10, Application US/10003806
; Publication No. US2002011929A1
; GENERAL INFORMATION:
; APPLICANT: Bishop, Colin E.
; APPLICANT: Agoulmik, Alexander I.
; APPLICANT: Zhu, Qichao
; TITLE OF INVENTION: CAN 1 AND ITS ROLE IN MAMMALIAN INFERTILITY
; FILE REFERENCE: P02066US1/10024824
; CURRENT APPLICATION NUMBER: US/10/003,806
; CURRENT FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: 60/245,872
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 173808
; TYPE: DNA
; ORGANISM: Mouse
US-10-003-806-10

Query Match 25.4%; Score 578.299; DB 1; Length 173808;
Best Local Similarity 63.2%; Pred No 1.5;
Matches 1147; Conservative 0; Mismatches 547; Indels 121; Gaps 18;
QY 292 TCTATTTCTTGATTTCTATCTTTGGGTCATTTTAACTCAGTAGTGAGTTGTTGGTTCC 351
Db 159106 TTTCTTTCTTTATCTCTTCCATCCAGAGATCATTCAGAGAGAGTTGTTGTTTCC 159165


```

; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 224
; LENGTH: 8048
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2005)..(2005)
; OTHER INFORMATION: n
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3503)..(3503)
; OTHER INFORMATION: n
US-10-388-934-224

Query Match      25.4%; Score 578.2; DB 1; Length 8048;
Best Local Similarity 62.0%; Pred. No. 5.8;
Matches 1160; Conservative 0; Mismatches 574; Indels 136; Gaps 17;

Qy      292  TCTATTTCTCATTTCTATCTGTGCTCATTTTAACTCAGTAGGAGTGTGTTGGTTTCC 351
Db      329  TTTCTTTCTTTATTTCTTTGACCGAGTATCATTTAGTAGAGCAATGTTCAATTTCC 3270

Qy      352  ATAAAGTTTGAAGTTTCTGTTGTTTCTGTTGTTTGTGTTTATCTAGATTAAAGCTGT 411
Db      3269  AGGTATATGTGGGCATTC-----TTCCTTATTTGTTATTGAACACCAAGTTTGGCCGT 3216

Qy      412  GGTGTCAGTAGGACATAGATATTTCAATTTGTTTATCTGTCGAGACTTCTT 471
Db      3215  GGTGTCGAGTAGCAGCATCGGATTTCTATCTTTCTGTTACCTGTGAGGCCGTTT 3156

Qy      472  TGTTTGAAATATGTTATCAATTTTGGAG--AGTTTCATAGGGTGTGTCACAAAGGTAC 529
Db      3155  TTTGACCAATATATGTTGTTCAATTTTGGAGAAAGTGCATAGAGGAGCTGAGAAGGTAT 3096

Qy      530  AGTCTTGTGTTTGGTGAATAGTCTGTAAA--TATCTCTAGTCTCACTTGTATTAGA 587
Db      3095  ATCCCTTTTGTGTTAGGATAGAAATGTTCTTAAATAATATCCGTTAAGTCCATTGGCTCATGA 3036

Qy      588  CATCAGTTAGTCCAGCATTTCTCTGTTTCGTTTGTGTTGAGATGACCTAACTGTTGGA 647
Db      3035  CTTCTCTTATGCTGTGACATCATCTGTTT--AATTTCTGTTTCCATGATCTGTCATTGAT 2977

Qy      648  GAGAAATGGGTATGAAATAGCCCACTATCTGTGTGTAGGT--CAATATGTGATTTTACG 706
Db      2976  GAGAGTGTGTGTTGAAATCTCCCACTATTTATTTGTGTGAGTGAATGTTGTTTTCAGC 2917

Qy      707  TGTAGCTGTGTTTATGAACTGTTGGTGCATTTGTTGTTGTCATAGACATTAAGA 766
Db      2916  TTTAGTAAGGTTTCTTTTACGTATGTAGTGCCCTTGTTTATTTGGGGCATAGATTATAGG 2857

Qy      767  ATTGCAATGTCTCTTGTGTGA--TTTTCTTTGAGCCCTATGTAGTATTTCTTCCCAATCT 825
Db      2856  ATTGAGAGTTTCTCTTGTGTGATTTTCTTTGATGAATGAATGAAGTGTCTTCTCTATCT 2797

Qy      826  CATCTGCTTATTTGGTTTAAAGTCTA--TTAGTCAGATATTAATTAAGTGTATCGCT 884
Db      2796  TTTTGTATGATCTTTTGTGTTGAAATGTTGTTTATTTGATATTTAGAAATGGCTTACTCCAGCT 2737

Qy      885  TGCCTTTTAGGGCCATTTGCTTAG--AATATCTTTTCCATCTTTTACTCTAAGGTGATGT 943
Db      2736  TGCCTTTCTGACCAATTTGCTTGGAAATTTGTTTTCAGCCCTTTTCACTCTGAGGTAGTGT 2677

```

```

Qy      944  CTATCCATG---GTAGTTGTCTTTTGGATGACAGTAGATGATGATGATCTTTTTCAT 1000
Db      2676  CTGTCTTTGTCTCTGAGGTGTCTTTCCITGAGCAGCAGAAATGCGGGTCTCTGTTGT 2617

Qy      1001  ATCCATCTCTTACCAGTATCTTTTCTAGAGAATTAAGATCATTTAGTCTATGATGT 1060
Db      2616  ATCCAGTTTGTAACTCATGACTTTTATTTGGGAGTTGAG-----GCCATGATGT 2565

Qy      1061  TGAGAATATCAATGAGCAGTGTTTTGTGATTTCTTTTATCTTTGCACTTTGTAAGTGT 1120
Db      2564  TGAGAGATATTAAGGAATAGTATTTGCTTTCCCGTTATATTTATTTGATGATGAGG 2505

Qy      1121  GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1180
Db      2504  TTATGTTTGTGTG----- 2492

Qy      1181  GTGTGTGTGTGTGTGTGTGTGTCTCTCCCTCTCTTTTGTATTTTGGCCCTGGAAATAT 1240
Db      2491  -----CTTTCACTCTCTTTTGTGTGCGCAAGCAGAT 2459

Qy      1241  TTATTAATCATATTTCTTGAATGTTGGGTAACAATCTTTAGATGAAAGTTTTCCTCTAGC 1300
Db      2458  TAGTTTCTTGTCTTCTAGGGTATAGCTTGCTCTCTATGTGGGCTTTTACCATTTAT 2399

Qy      1301  CTCTTT--TAGGTCTGCATTTGAAGATAGATATTTCTTACATCTGATTTTATCTTAGAAT 1358
Db      2398  ATCTTTGTAGTCTGGAATTTGTGGAAGATATTTGTGTAATTTGGTTTGTGCAATGGAAT 2339

Qy      1359  GTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCTGA 1418
Db      2338  GTCTTGGTTTCTCCATCAATGTTAATTGAGAGTTTGTCTGGATACAGTAACCTGGGCTGG 2279

Qy      1419  CATCTGTAGTCTCTGGAGTCTGTAGCACAATCTGTGCAGGGCTTCTTACATTTTGAAT 1478
Db      2278  CATTTGTGTCTCTTAGGGTCTGTATAACATCAGTCCAGGATCTTCTGGCCCTCATAGTT 2219

Qy      1479  TCTATTGGAAAAGTCAAGGTGTAATCTAATAATCTGCTCTTATATGTTAAATGGTCTTT 1538
Db      2218  TCT--GCGGAGAAGCTGTGGTGTGATCTGATAGTCTCCCTTTTATATGTTACTT--GACCTT 2161

Qy      1539  TTTCCCTTGCATCTTTTAAATATCTTTCTTGTCTATATCTTTTACTGATGATTTATTA 1598
Db      2160  TTTCCCTTACTGTTTTAAATATCTTTCTTTTATTTTGTGCGTTTGGTGTGTTTGAACAATTA 2101

Qy      1599  TGCACGTGTGGGAGTTTCTTTCCGGTCCAACTATTTGGTGTGTTTGTATGTTCTTTGTA 1658
Db      2100  TGTGACGGGAGGTGTTTCTTTTCTGTCCAACTATTTGGAGTTCTGTAGGCTTCTTGTG 2041

Qy      1659  CTTGTATAGGCATCTCTTTCTCAAGGTAGGAAATTTTCTTTTGTGTTTCTTTGAAAA 1718
Db      2040  TGTCTATGGGTATCTCTTTTGTAGTTAGGGAAGNTTCTTCTATGTTTGTGGAAGA 1981

Qy      1719  TATTTTCCCTGCTTTTGTGCTG-----CCTTCTTCCCTTCTCTATCTCT 1764
Db      1980  TATTTACTGCTTCTTGTAGCTGGAGTCTTCACTCTCTTCTATACCTATTTATCTTAGGT 1921

Qy      1765  TTGGTTTTTGCATAGTGTCTCTGGCTTCTCGGATGTTTATGCTGGAATTTATTTAGACT 1824
Db      1920  TTGATCTTCTCATGTAGTCTCGAATTTCTGTATGTTTGGACCAAGTACTTTTCCGCT 1861

Qy      1825  TAAATTTCTTTGACCAAGGATATCCATTTCTTATCTGTCTTCACTGCTGAGATTC 1884
Db      1860  TTAATATCTTTGACAGTTGAGTCAATGATTTCTATGGAATCTTCTGCTGCTGAGATTC 1801

Qy      1885  TCTCTTCTATCTTGTATTTCTGTGAGAGCTTGTCTCTGAGGTCTCTGTTGGTGTCT 1944
Db      1800  TCTCTTCCATCTCTTGTATTTGTGTGGAAGCTTTTATCTACAGCTCTCTGCTCTCT 1741

Qy      1945  TAAATTTTTC--AATTTCCAGATTTCTTCAAGTTTGGTGTGTTTGTATTAATTTCTATTTCA 2003
Db      1740  TTTGGTTTTCTATATCCAGGTTGTTTCCATGTGTTCTTTTGTGTTGTTGTTCTATTTCA 1681

```



```
QY 885 TGCTTCTTAGGCCATTTGCTTAG-AATATCTTTTCCATCTTTTACTCTAAGGTGATGT 943
Db 2736 TGCCTCTTCTGACCAATTTGCTTGGAAATTTGTTTCCAGCTTTTCACTCTGAGGTAGTGT 2677
QY 944 CTATCCATG---GTAGTGTGCTTTTGGATGAGCAGTAGGATGAGTCTTGTTCAT 1000
Db 2676 CTGCTTTGCTCTGAGGTGTTTCTGTTAGGCAGCAGATGAGGCTCTCGTTGCT 2617
QY 1001 ATCAATCTGTATCCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCATGATGT 1060
Db 2616 ATCCAGTTTGTAACTATGACCTTTTATTTGGGAGTTGAG-----GCCATTTGATGT 2565
QY 1061 TGAGAAATATCAATGAGCAGGTGTTGGAATCTTTGTTATCTTGCACCTTTGGAAGTGTGT 1120
Db 2564 TGAGAGATATTAAGGAATAGTATATGCTTCCGTTATATTCATATTTGGATGTGAGG 2505
QY 1121 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1180
Db 2504 TTAATGTTTGTGTG----- 2492
QY 1181 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1240
Db 2491 -----CTTTCATCTCTTTGTTTGTGTTGTTGCTGGAATAT 1240
QY 1241 TTAATTAATCATATTTTCTGAATGTGGGTAAACATCTTTTAGATGAAGTTTTCCTTAGC 1300
Db 2458 TAGTTTCTGTCTTCTTAGGATATAGCTTCCCTCTTATGTTGGGCTTTACCAATTTAT 2399
QY 1301 CTTCTT--TAGTCTGATTTGAGATAGATATCTTTTACATCTGATTTATCTTGAAT 1358
Db 2398 ATCCCTTTGTAGTGTGATTTGTGGAAGATATTTGTGAAATTTGTTTGTGATGAAT 2339
QY 1359 GTCTTTCTTTCTCCAACTATTTGTGACAGAAATTTTCTAAGTGCAGTAGTCTGGCTGA 1418
Db 2338 GTCTGTGTTTCTCCATCAATGTTAATGAGAGTTTGTGATGATACAGTAACCTGGCTGG 2279
QY 1419 CATCTGTAGTCTTTGGAGTCTGTAGACACATCTGTGAGGGCTTCTTACATTTTGATTT 1478
Db 2278 CATTTGTGTTCTTTAGGCTCTGTATAACATCAGTCCAGGATCTCTCTGGCTTCATAGTT 2219
QY 1479 TCTATTGGAAGAGTCAGGTGAATTTCTAATCATCTGCTTTATATGTTAATTTGGCTTT 1538
Db 2218 TCT-GCCGAGAGAGTCGTGTGATCTGTATAGTCTCCCTTTATATGTTACTT--GACCTT 2161
QY 1539 TTCCCTTTGCATCTTTTAAATATCTTTCTTTGTCTATATCTTTTAGTATGATTTGATTA 1598
Db 2160 TTTCCTTTTACCTTTTAAATATCTTTCTTTATTTTGTGCTTTTGGTGTGTTTGACAA 2101
QY 1599 TGCACTGTGGGAGTTCTTTCCGGTCCAACTATTTGTTGTTTGTATGCTCTTGTGA 1658
Db 2100 TGTGACGGGAGGTGTTCTTTCTGCTCAATCTATTTTGGAGTTCTGTAGGCTCTTTGTG 2041
QY 1659 CTTGATAGGCATCTCTTTCTCAAGGTTAGGAAATTTTCTTTTGTGTTTCTTGAAAA 1718
Db 2040 TGTCTATGGATATCTTTTATAGTTTAGGAGAGTTTCTTCTATGTTTGTGTTGAAGA 1981
QY 1719 TATTTTCCCTGCTTTGACCTG-----CCTTCTCCCTTCTCTATTCCT 1764
Db 1980 TATTTTACTGGTCTTTGAGCTGGGAGTCTTCACTCTCTCTATACCTATTAATCTTAGT 1921
QY 1765 TTGGTTTTTGTAGTGTCTCTGGCTCTCGATGTTTATATGCTCGGATATTTTAGACT 1824
Db 1920 TTGATCTCTCATGAGTCTCGATTTCTGTATGTTTGGACCAAGTAGCTTTTTCGCT 1861
QY 1825 TAAATTTCTTTGACCAAGGATACATTTCTTCTATCTTGTCTTCACTGCTGAGATTC 1884
Db 1860 TTACATATCTTTGACAGTTGAGTCAATGATTTCTATGSAATCTTCTGCTCCTGAGATTC 1801
QY 1885 TCTCTTCTATCTTGTATCTGTCTGAGGAGCTTTCTCTGAGGTTCTCTGCTGCTGCTTCT 1944
Db 1800 TCTCTTCCATCTTGTATCTGTGTTGAGAGCTTTTATCTACAGCTCTTGTCTCTTCT 1741
QY 1945 TAAATTTTTTCT-ATTTCCAGATTTCTCTCAGTTTGGGTTTGTGTTAATTAATTTTCCA 2003
```

```
Db 1740 TTTGGTTTCTATATCCAGGTTGTTTCCATGTGTTCTTTTGTGATGCTTCTATTCCA 1681
QY 2004 CTTTCAGGTCTGGAATGTTTACTCATTTCCCTCCAGTATTTACATTTTTCATAGGTTT 2063
Db 1680 TTTTAAATCTCTCAACTGTTT-----GATTTGTTTCTCTGGAATTC 1638
QY 2064 CTTTAAATGATTTATTCATTTCTCTTCAAGGACCTTTTATGAATTCATAAAATGTATGT 2123
Db 1637 CTTTCAGGATTTTGTGTCTCTCTATGGCTTCTACTTGTGTTTATTTATGTTTCT 1578
QY 2124 TAAGGTCTTT 2133
Db 1577 GGAATTCCTTT 1568

RESULT 25
US-10-087-192-547/c
; Sequence 547, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 547
; LENGTH: 116858
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(116858)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-547
```

```
Query Match 25.2%; Score 574.999; DB 1; Length 116858;
Best Local Similarity 62.1%; Pred. No. 1.8; Gaps 18;
Matches 1150; Conservative 0; Mismatches 555; Indels 124; Gaps 18;

QY 292 TCTATTCTTGATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGAGTTGTTGTTTCC 351
Db 37530 TTTCTTTCTTTTCTTTCTTTCTTTGATCAAGGTATCATTTGATAGAGTTGTTGTTTCC 37471
QY 352 ATAAGTTTGTAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 411
Db 37470 AAGTGAATCTGGCTTTCTATTATTATGCTGTTATTGAAGAT---CAGCCTTAGTCCGT 37414
QY 412 GGTGGTCAGATAGGACATAGATATTATTCAATTTGCTTTTATCTGTCGAGACTTGCTT 471
Db 37413 GGTGATCTGATAGGATGATGAATATTTTAAATTTTGTATCTGTTGAGCCTGTTT 37354
QY 472 TGTTTTGAATATGTATTCATTTTGGAG--AGTTTCATAGGGTCTGTACAGAGAGGTAC 529
Db 37353 TGTGATCAATATATGTTCAATTTTGGAGAGGTACCATGAGGTGCTGTAGAGAGAGGTAT 37294
QY 530 AGTC-TTTCTGTTTGTGTAATAGTCTGTAATAATCT-CTAGGTCCACTTGGTTTATGA 587
Db 37293 ATCTTTTGTTTTGTAGATATAATGTTTGTAGATATCTGTTAGATCCATTTGTTTCAAA 37234
QY 588 CATCAGTTAGCTCCAGCATTTCTCTGTTTCTGTTTCTGTTTGTGAGATGACCTAACTGTGA 647
Db 37233 CTTCTGTAGTTTCACTGTGCTGCTGTTTAG-TTCTGTTTCCAGGATCTGTCCATGTTGT 37175
QY 648 GAGAAATGGGTTATGAGTAGTACCCACTATCTGTGTGAGGT-CAATATGTGATTTTACG 706
```

Db 37174 GAAAGTGGGTGCTGAAAGTCTCCCACTATATATGTTGGAAGTCAATATGTTGCTTTGAGC 37115
Qy 707 TGTAGCTGTGCTGTTTATGAACCTGGGTGACATTTGTTTGGTGCATAGACATTAAGA 766
Db 37114 TTTACTAAAGTTTCTTTAATGAATGCTGCTGCAITTTGCAITTTGGAGCATATATTCAGA 37055
Qy 767 ATTGCAATGCTCTTGTGTGAATTTT---CCTTGAAGCTCTAGTAGTATTTCTCCCAATC 824
Db 37054 ATTGAGAGTTTCTCTGTGTAGATTTTACCTTTTGATGAATTAAGAGTGTCTTGTCTTTT 36995
Qy 825 TCATCTGCTTATGTTTGGGTTTAAAGTCTATTAGTCAAGATATTAATAAGTCACTGTATCGCT 884
Db 36994 TTGATAACTTTGGGTTGGAAGTCAAT---TTTATTTGATTTAGAAAGTCTTCCAGCT 36938
Qy 895 TGTCTTTAGGCGCAATTTGCTTAGAATA-TCTTTCCATCTCTTTTACTCTTAAGTGAATG 943
Db 36937 TGTCTTTCTATACCAATTTGCTTTGGAATAATGTTTCCAGCCCTTTTACTCTGAGTAGTGT 36878
Qy 944 CTATCCAT---GGTAGGTTGCTTTTGGATGAGAGTAGTAGGATCTTGTGTTTTCAT 1000
Db 36877 CTGCTTTTATCCCTGAGTGGGTTTCTGTAAAGCAGCAAAATGTTGGTCTCTGTTGTGT 36818
Qy 1001 ATCCATCTGTATCCCAATATCTTTTCTAGAGAAATTAAGATCAITTAGTCAATGATGT 1060
Db 36817 ACCAGTTGTTAGTCTATGCTTTTCTTGGGAAATTGAG-----TCCATTTGATGT 36766
Qy 1061 TGAGAAATTAATCAATGAGCAGTGTGTTGGATCTTGTATCTTGCATCTGTGAAGTGTGT 1120
Db 36765 TAAGAGAAATCAAGAGAAAGTAATGTTGTTTCTGTCTCAITTTTGTGTTAAAGTTGGGA 36706
Qy 1121 GT 1180
Db 36705 ATCTGTTCTTATGCTGTCTTCTTTAGTTTG----- 36673
Qy 1181 GTGTCTGT 1240
Db 36672 -----TTAAAGGAT 36664
Qy 1241 TTATTTATCATATTTCTTGAATGTTGGTAACTTTTAGATGAAGTCTTCTCCCTAGC 1300
Db 36663 AACCTTCTGCTTTTCTAGGTTGAATTTCCATCTTGTGTAGTGTGTTTCCCTTTAT 36604
Qy 1301 CTTCTTT---AGGCTGCAATTTGAAGATAGATATTTTACATCTGATTTTATCTTAGAAT 1358
Db 36603 ATCTTTGAAGGCTGGATTCGTGAAAGATATTCGGTGAATTTGGTTTGTCTCATGAAT 36544
Qy 1359 GTCTTTCTTTTCCCACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCTGA 1418
Db 36543 ACTTTGATTTCTCCATCTATGTTAATTGAGAGTTTGGCTGGGTATAGTCTGGCTGG 36484
Qy 1419 CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGCTCTTACATTTTGAAT 1478
Db 36483 AATTGTGTTCTCTTGTGTCTGTATACATCTGTCCAGGATCTTCTGGCTTTCATAGTC 36424
Qy 1479 TCTATTGAAAGTCAAGTGTAAATCTTAATACATCTGCTTTTATATGTTAATGCTCTTT 1538
Db 36423 TCT---GGAGAGAAGTCTGGAGTAATTTCTAATAGGCTGCTTTATATGTTACTT---GACCTT 36366
Qy 1539 TTTCCCTGTGCATCTTTTAATATCTTTCTTTGTTCTATCTTCTTATAGTATTTGATTA 1598
Db 36365 TTTCCCTTACTGCTTATAATATCTATCTTTTATTATTAGTCAATTTATTGTTCTGAATTA 36306
Qy 1599 TGCAGTGTGGGAGTTTCTTTTCCGGTCCAAATCTATTGTTGTTTGTATGCTTCTTGTA 1658
Db 36305 TGTCTGGGAGGAATTTCTTTTCTGGTCCAGTCTATTGGAGTACTGTAGGCTTCTTATA 36246
Qy 1659 CCTTGATAGGCATCTCTTTCTCAAGTTAGGAATTTTCTTTTGTGTTTCTTGAAAA 1718
Db 36245 TGTTCATGGGCATCTCTTCTTTTAGGTATGGAAATTTTCTTCTATATATTTTGTGAAGA 36186
Qy 1719 TATTTTCCCTGCTTTTGACCT-----GCCCTCTCCCTTCTCTATTCCTTT 1766

Db 36185 TATTGCTGGCCCTTTAAGTTAAATCTTCATCTCACTACTCTATTAATCCGTAGTTT 36126
Qy 1767 GGTTTTTCATAGTGTCTCTGGCTTCTCGATCTTTTATGCTCGGATTTATTTAGACTTA 1826
Db 36125 GGTCTTCTCATCTGTGCTCGATTTCTCGATGTTTGTAGTAGGATCTTTTGCATTT 36066
Qy 1827 ACATTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTGAGATCTC 1886
Db 36065 ACATTTCTTTGATGATATATCCATGCTCCCTATGGAATCTTCTGACCTGAGATCTC 36006
Qy 1887 TCTTCTATCTCTTGTATTTCTGTCAGTGAGCTTGTCTGAGGTTCTG-TTGGGTTCTT 1945
Db 36005 TCTTCCATCTCTTGTATTTCTGTGCTGATGCTCAGATCTATGTTCTGATTTCTTCT 35946
Qy 1946 AATTTTTCATTTCCAGATTTCCCTCAGTTTGGGTTTCTTTTATTAATTTCTATTTCCACT 2005
Db 35945 AGGTTTCTATCTCCAGATTTGCTTCTTGGGTTTCTTCTATTTCTTCTTCTTCTATT 35886
Qy 2006 TTCCAGTCTGTAATCTTTTACTCAATTTTCTCT---CCCAGTATTTACATTTTCATAGTT 2062
Db 35885 TTTAGATCTTGGATGTTTTTGTTCATTCATTCGCTGTTGTTGTTTCTTCTGTTTCTCTAAT 35826
Qy 2063 TCTTTAATGATTTTATTCATTTCTCTTCTC 2091
Db 35825 TCTTTAAGGATTTTGTGTTTCTTCTTCTAC 35797

RESULT 26

US-10-087-192-1033
; Sequence 1033, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1033
; LENGTH: 200400
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)._(200400)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1033

Query Match 25.2%; Score 573.4; DB 1; Length 200400;
Best Local Similarity 63.4%; Pred. No. 1.5;
Matches 1164; Conservative 0; Mismatches 546; Indels 126; Gaps 20;
Qy 292 TCTATTTCTTGTATTTCTATCTTGGCTCATTTTAACTCAGTAGTGTGTTGTTTCC 351
Db 157315 TTCTTTCTTAATTTCTTACCTGACTAATTT-----AGAGTAGAGAGTTGTTTCAGTTTCC 157369
Qy 352 ATAAGTTTCTAAGTTTCTGTTGTTTCTGTTGTTGTTTATCTAGATTTAAGCTGT 411
Db 157370 ACCTGATCTGGCTTTTGTGCTTTTGTCTGTTTATTAAGA---CCAGCTTTAAGCCAT 157426
Qy 412 GGTGTGCATAGGACATAGATATTATTCAATTTGTCTTTTATCTGTGCGAGCTTGT 471
Db 157427 GGTGATCTCATAGGATGATGGGATTTATCTGCTCTTCTGTATCTGTTTAGGTTGTT 157486
Qy 472 TGTTTTGAATATGATTTCAATTTTGGAG--AGTTTTCATAGGCTGCTGACAGAGGTAC 529

157487 TGTGACCAATTATATGTCAGTTTGGAGAAAGTACCATGAGTGTCTAAGAGAAGCTGT 157546
QY 530 AGTC-TTTGTGTTTGGTGAATAGTCTGTAAATATCT-CTAGTCCACATTTGGTTTATGA 587
DB 157547 ATTCTTTTGGTTAGGGTGAATAGTCTATAGATATCTGTAAATCAATTTGGTTCAATA 157606
QY 588 CATCAGTTAGTCCAGCAATTTCTCTGTTTCCTTTTGGTGGAGAGCACTAACTGTTTGA 647
DB 157607 CCTCTATTAGTTTCACTGTCTCT-TTTTATTCTGTTTCAATGACCTGTCCATTTGAT 157665
QY 648 GAGATGGGTATTGAAGTACCCACATCTGTGTGT-GAGGTCAATATGTGATTTTATGC 706
DB 157666 GACAGTGGGTGTTGAGTCTCCCACTATTAATGTGTAGGGTTCATGTGTGTTTGGC 157725
QY 707 TGTAGCTGTCTGTTTATGAACCTTGGTGACATTTGTTGTTGGTGCATAGACATTAAGA 766
DB 157726 TTTAGTAAAGTTTCTTTTATGAATGTGGGCGCTCTTGCAATTTGGGCGATAGATTTCAA 157785
QY 767 ATTG-CAATGCTCTTGGTGGATTTT-CTTTGATGCTATGTAGTATTTCTTCCCAATC 824
DB 157786 ATTGAGATTTTATCTTGATGATTTTCCCTTGATGAATAGGAGTGTCTCTCTCATC 157845
QY 825 TCATCTGCTTGTGTTTGGTGTAAAGTCTA-TTAGTCAATATTAATGAATGACTGTATCGGC 883
DB 157846 AGCTTGTATTAATTTGTTGTAAGTCTATTTATTTGATATAGGATGACTCCAAAC 157905
QY 884 TTGCTTCTTAGGGCAATTTGCTTGAATATCTTTTCCATCTCTTATCTCTAAGTGTAT 943
DB 157906 TTGTTTCTGTGACCAATTTGCTTGAAGACCTTTTCCATCTCTTATCTCCGAGATGAT 157965
QY 944 CTATCCATGTA-GGTTGCTTTTGTGATGATGAGTGTGATGATGATGATGATGATGAT 1002
DB 157966 CTGCTTTGTTATGAGGTGTCTTATATGCAAAAGTCTGGGTTTGTGTGTAT 158025
QY 1003 CCATCTCTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATAGTGAATGATGTTG 1062
DB 158026 CCAGTCTGTAGTCTATGCTTTTATAGGTGAGTTGAG-----TCTATTAAATATG 158077
QY 1063 AGAATATCAATGACGAGTGTGAGTCTGTTGATCTTGTATCTTGCATCTGGAAGTGTGT 1122
DB 158078 AAAGATATAAGAGAAAGTGTGTTGTTCTCTG----- 158111
QY 1123 GGT 1182
DB 158112 -----ATGTTGTTTCAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 158151
QY 1183 GTCGTGTCTGT 1242
DB 158152 -----CTCCTTATGACTTTGTTGTGAGATGCTTA 158180
QY 1243 ATTATTCATATTTCTTGAATGTGGCTAACTTTAGATGAAATTTTCTCTAG--C 1300
DB 158181 ATATCTTGTCTTTCTAGTGCAGATATCTCTGTGTATTTAGTTTTCCTCTAGAT 158240
QY 1301 CTTCTTTAGGTGCTGATTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGATGT 1360
DB 158241 CCTCTGAAGGGCTGATTTGGTATAGATATCTGTTTGAATTTGGTTTGTCTCTGGAATAT 158300
QY 1361 CTTTCTTTCTCAACTATTGTGACAGAAAGTTTCTTAAGTGCAGTGTGTGGCTGACA 1420
DB 158301 TTTGTGTTTCCCCATATGTTGATTGAGAGTTTGTCTGGGTGTAGTAGCCTAGCTGCA 158360
QY 1421 TCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGAGGGCCCTTCTTACATTTTCAAGTTTC 1480
DB 158361 TTTGTGTTCTCTTAGAGTCTGTATGACCTCTGGCCAGGCTCTTCTGGCTTTTCAATGCT 158420
QY 1481 TATTGGAAAGTCAAGTGTATTTCTAATACATCTGCTTTATATGTTAAATGGTCTTTT 1540
DB 158421 TGTT-GAGAAGTCTGATGATTTCTGATAGGTGCTGCTTTGTATGTACTTTGG-CCATTT 158478
QY 1541 TCCCTTGTACATCTTTTAAATATCTTTTCTTGTGTCTATCTTTTGTAGTGAATTTATTTATG 1600
DB 158479 TCCCTTGCAGCTTTTAAATATCTCTCTTTGTTCTGTGGTGTGTGTGTGTGTGTGTGTGT 158538

QY 1601 CACTGTGGGAGTTTCTTTTCCGGTCCCAATCTATTGGTGTGTTTGTATGCTTCTGTGACC 1660
DB 158539 TCACAAGAGCATTTTCTTTTCTGGTCCCATTTATTGGTGTGTTTCTATATGCTTCTGTGACC 158598
QY 1661 TTGATAGGCACTCTTTCTCAGAGTTAGAAATTTTCTTTTGGTGTGTTTCTTGTGAAATA 1720
DB 158599 TTTATGGCCACTTCTTTCTCAGGTGGGAAAGTTTCTCTATGATTTTGTGTGAGACA 158658
QY 1721 TTTTCCCTCTTTTGACCTG-----CCTTCTTCCCTTCTCTATTTCTTTT 1766
DB 158659 TTTTATAGTCCCTTTGAGATGGGAACCTTCAATCTCTCTATTTCCAAATATTCTTAGATTT 158718
QY 1767 GGTGTTTGCATAGTCTCTGCTTCCCTGAGTGTGTTTATGCCCTGGATTTATTTAGACTTA 1826
DB 158719 GGTCTTTGATGTGTCCTGAATTTCTCTGGATGCTTGGGT-----AGTTTGTACATTT 158774
QY 1827 ACATTTTCTTTGACCAAGATGATCCATTTCTCTATCTTCTCTACCTGCTGAGATTTCTC 1886
DB 158775 GAATTTTCTTGCAGAGTTGTGTGAGTCTCTCTACTGTATCTTACACCTGAGATTTCTC 158834
QY 1887 TCTTCTATCTTGTATTTCTGTCAGTGAAGCTGTCTCTGAGGTTCTCTG-TTGGGTTCTT 1945
DB 158835 TCTTCTATCTTGTATTTCTGTGTGATGCTTACATCTGTAAATCTCTGACCTCTCTCT 158894
QY 1946 AATTTTCTTATTTCCAGATTTCTTCTCAGTTTGGGTTTGTGTTTATTAATTTCTATTTCCACT 2005
DB 158895 AGTTTTCCTTTTCTCAGTTTGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 158954
QY 2006 TTAGTCTCTGAAATGTTTACTCAATTTCTCTCCAGTATTTACATTTTCTAGGTTTCT 2065
DB 158955 TTTACTTGTATCGCTTTATTAATTTCTCTCACTGTTTACCTGTGTTTCTCTGTTTCT 159014
QY 2066 TTAATGATTTTATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2101
DB 159015 TTAGTGTGATTTAATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 159050

RESULT 27

US-10-175-523-57
; Sequence 57, Application US/10175523
; Publication No. US20030095264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/13795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57
; LENGTH: 225883
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-57

173126	CC	TTT	TGAAGGGCTGGAATTCGTGGAAAGATATATGTGNAATTTGGTTTGTTCGTGGAATAC	173185
1361	CTT	TC	TTTTCCTCCAACTAATGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGCCCTGACA	1420
173186	TTT	CG	TTTCTACATCTATGGTAAATTGAGAGTAGGC-ACGTATAGTAGCCTGGGCTGGCA	173244
1421	TC	GT	AGTCTCTTGGAGTCTGTACACATCTGTGCAGGGCCTTCTTACATTTTGTAGTTTC	1480
173245	TTT	GT	GTCTCTTAGTGTCTGTATAACAATCTGTCCAGGCTCTTCTGGCTTTCATAGTCTC	173304
1481	TAT	PG	AAAAAGTCAGGTGTAATTTCTAATACATCTGCCCTTATATGTTAATTTGGTCTTTTT	1540
173305	TGG	T	-GAAAAGTCTGGTGTAAATCT-ATAGGCCCTGCCCTTATATGTTACTT-GACCTTTC	173361
1541	TC	CC	TTCATCTTTTAAATATCTCTTCTTTGTCTATACCTTTTAGTGAATTCATATATTATG	1600
173362	TCC	CT	TACTGCTTTTAAATATCTCTCTTTATTTAGTGCATTTGTTGTTCTGATTAATTATG	173421
1601	CAC	TG	TGGGAGTTCTTTTCCGGTCCAACTATATTTGGTGTTTGTATGCTTCTTTGTACC	1660
173422	TGT	CG	GAGGAAATTTCTTTTCTGGTCCAGTCTATTTGGAGTCTCTGTAGGCTTCTTTGTATG	173481
1661	TTG	AT	GAGCATCTCTTTCTCAAGGTTAGGAATTTTCTTTTCTTTTGGTTTCTCTGAAATA	1720
173482	TT	CA	TGGCATTCTCTTTAGGTTTGGGAAGTTTCTCTTAATTTTGTGGAATATA	173541
1721	TTTT	CC	CTGCTTTTGACCTG-----CCTTCTTCCCTCTCCTCTATTCCTTT	1766
173542	TTT	GC	TGGCCCTTTTAAAGTTGAAAATCTTCATCTCTCAACTCTTATTCGTAGGTTT	173601
1767	GGT	TT	TGTCATCTCTGATCTCTGGCTTCTCGAAGTTTATGCTCGAATTTATTAGACTTA	1826
173602	GGT	CT	CTCTCATGTGCTTAGATTTCCGTGGAGTTTTCAGTTAGATCTCTTTTGCAATTT	173661
1827	ACA	TT	TTCTTTTGACCAAGGATCCATTTTCTATCTTGTCTTCTCACTGCCGTGAGATTCTC	1886
173662	GTA	TAT	CTTTGATTTGTTGTGCCGATGTTCTCTATGGAACTCTTCTGCCCTTGAGATTCTC	173721
1887	TC	TT	CTATCTCTGATATCTCTGTCAGTAGGCTTGCTCTGAGGTTCCGTG-TTGGGCTTCTT	1945
173722	TC	TT	CCATCTCTGTATCTGTCTGCTGATCTCTGTCTATGGTCTCCAGATTTCTTCTCT	173781
1946	AA	TT	TTTTTTTCATTTCCAGATTTCTCTCAGTTTGGGTTTTGTTTTATTAATCTATTTCCACT	2005
173782	AGG	GT	TTCTATCTCCAGTGTGCTCCTCACTTTGGGTTTCTTTTATTGTGTCFACTTCCCTT	173841
2006	TT	CA	GGTCTCTGAAATGTTTTACTCATTTTTCTCCGAGTATTTACATTTTTCATAGGTTCT	2065
173842	TTT	AG	GTCTAGATGGTTTGTTCATTTCCATCCCTGTTTGGATGTGTTTTCCCTGTTTT	173901
2066	TTA	AT	TGGATTTATTTCAUATTCCTCTTCAAGGACCTTTTATGAATTCATAAAATGATGTTA	2125
173902	TCT	AT	AAGACTTGTAACTCTTTAGCAGTGTTCCCTGTAATTTCTTTAAGTGAGTTATTA	173961
2126	AGG	TC	CTTT	2133
173962	AAG	TT	CTTT	173969

RESULT 28
US-10-087-192-2029
; Sequence 2029, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01

Db 159422 TCTGTTCTTGGCGTCTCTCTTTAGGTTGTGTAGAGATTACCTCTCTGTTTTTCT 159481
Qy 1181 GTGCTGTGTCTGTGTGTGTCTCTCTCCCTCTTTTGTGATTTTGGCTGGAATTAT 1240
Db 159482 AGGGCAATGTTCCCGT----- 159498
Qy 1241 TTAATTAATCAATTTCTTTGAATGTGGTAACAATCTTTAGATTGAAGTTTTTCTCTAGC 1300
Db 159499 -----TACTGTATTGGTTTTTCTCTGTTATT 159525
Qy 1301 CTTCTTTT--AGGCTGCAATTTGAGATAGATATCTTTTACATCTGATTTATCTTAGAAT 1358
Db 159526 ATCTCTTGAAGGCTGGATTCTGTGAGAGATAATGTGGAATTTGGTTTTGTCTGGAAT 159585
Qy 1359 GTCTTCTTTCTCCA-ACATTTGTGACAGAAAGTTTTTCTAAGTGCAGTAGCTCTGGCTG 1417
Db 159586 ACTTTGGTTTCTCCATCTATGGTAATTGAGATTGGCTGGGTATAGTAGCTGGCTG 159645
Qy 1418 ACATCTGTAGTCTCTGTGAGTCTGTAGACATCTGTGAGGCGCTCTTACATTTGAGT 1477
Db 159646 GAATTTGGTTCTCTTAGTGTGTATACATCTGTCCAGGCTCTCTGGCTTTTCATAGT 159705
Qy 1478 TCTATTGGAAAGTCAGGTGTAATCTTAATACATCTGCCCTTATATGTTAATTTGGTCTT 1537
Db 159706 CTCTGGTGAATAA-TCCTGTGTATCTGATGCTGCTTGGCTTTGTATGTTACTT-GACCT 159763
Qy 1538 TTTCCCTTGCATCTTTTAAATCTTTCTTTCTTTGTTGTTATATCTTTTGTGATTTGATTT 1597
Db 159764 TTTTCCCTTACTGCTTTTGTATCTATCTTTTATTTAGTGCATTTGTTGTTCTGATTT 159823
Qy 1598 ATGCACTGTGGGAGTTCTTTTCCGGTCCCAATCTATTTGGTGTGTTGTATGCTCTTGT 1657
Db 159824 ATGTCTGGGAGGAGTTCTTTCTGGTCCAGTCTATTTGGAGTCTGTAGGCTCTTGT 159883
Qy 1658 ACCTTGATGAGCATCTCTTCTCAAGGTAGGAAATTTTCTTTTGGTTTTCTTGAAA 1717
Db 159884 ATGTTCAATGATCTCTCTTTCTTATATTTGGGAAGTTTCTTCAATAAATTTTGTGAAG 159943
Qy 1718 ATATTTTCCCTGCTTTGACCTG-----CCTCTTCCCTGCTCTATCTCC 1763
Db 159944 ATGTTGTGCTGCTTTGAGTTGAAATCTTCAATCTCATCACTCTCTATATCCGTAGG 160003
Qy 1764 TTTGGTTTTTGCATAGTCTCTGGCTCCCTGGATGTTTATGCTGGATTTATTTAGAC 1823
Db 160004 TTTGGTCTTCTCATGTGTCTGCTGGATTTCTGGATATTTTGTAGTAGATCTTTTGCAT 160063
Qy 1824 TTAACATTTTCTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCCCTGAGATT 1883
Db 160064 TTTCCATTTTCTTGAATGTTGTGCGCATGTTCTCTATGGAATCTTCTGCACCTGAGATT 160123
Qy 1884 CTCTCTTCTATCTCTGTGATCTGTGAGTGGCTTGTCTGAGGTTCCCTG-TTGGGTT 1942
Db 160124 CTCTCTTCCATCTCTGTATCTGTGCTGATGCTCAAAATCTATGTTCCAGATTTCTTT 160183
Qy 1943 CTTAATTTTTTCAATTCAGATTTCTTCAAGTTTGGGTTTGGTTTATTAATCTCAATTC 2002
Db 160184 CCTAGGTTTCTATCTCCAGTGTGCTCACTTTTGGTTTCTTTTATTTGTCTACTTCC 160243
Qy 2003 ACTTTCAGTCTGAAATGTTTACTATTTTCTCTCC-AGTATTTACATTTTCTAG 2059
Db 160244 CTTTATAGGTAGTAGTGGTTTGTTCATTTCCATCCACCTGTTTGTATGTTTTTCTCT 160303
Qy 2060 GTTCTTTTAAATGGATTTTATTCATTTCTCTTCAAGGACCTTTTATGAATTCATAAATGT 2119
Db 160304 TTTTCTAGGACTTCTACCTGTTTGTATGTTTCTTCTTTTAAAGACTTGT 160363
Qy 2120 A 2120
Db 160364 A 160364

RESULT 33

US-10-374-077-209/c

; Sequence 209, Application US/10374077
; Publication No. US20040006779A1
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; Yu, Chang-En
; Oshima, Junko
; Mulligan, John T.
; Schellenberg, Gerald D.
; TITLE OF INVENTION: ANTIBODIES AGAINST GENE PRODUCTS RELATED TO
; WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/10/374,077
; FILING DATE: 25-Feb-2003
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman, Stephen
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 100107.401D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 209:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51259 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 209:
US-10-374-077-209

Query Match 24.8%; Score 565.399; DB 1; Length 51259;
Best Local Similarity 62.1%; Pred. No. 2.9;
Matches 1155; Conservative 0; Mismatches 576; Indels 128; Gaps 18;
Qy 292 TCTATTCTTGTATTTCTATCTTGGCTCATTTTAACTCAGTAGTAGTGTGTTGTTTCC 351
Db 10412 TCTTCATTTTCAATTCCTGCTTGACCAAGGTATCATTTAGTAGAGCGTGTTCAGCTCC 10353
Qy 352 ATAGAGTTTGTAGTTTCTGTGTTCTGTTCTGTTGTTGTTGTTGTTATCTAGATTAACTGT 411
Db 10352 ATATGATGTGTGCTTCCGTTGTTGTTGTTTAAAGACCAATCTTAGGCCGTGTT 10293
Qy 412 GGTGGTCAGATAGGACATAGATTATTTCAATTTGTTCTTTTCTGTCGAGACTTGTCT 471
Db 10292 GGTCTCATGATAGGTGTCATGATGATTTCCATCATCTTGAATCTGTAGAGTCTGTT 10233
Qy 472 TGTTTTGAATAATGATTTCAATTTTGGAGA--GTTTCATAGGTGTCGACAGAGGTAC 529
Db 10232 TGTGACCACTATATGTCAGTTTGGAGAGGTTCCATGAGGTGCTGAGAGAGAGTAT 10173
Qy 530 AGTCTTTGTGTTTGGTGAATAGTCTGTAATATCT-CTAGTCCACTTGGTTTATGAC 588
Db 10172 ATTATGTTTGGATGACATGTTCTTATAATATCTGTTAGATCATTTGGTTTCAATAC 10113
Qy 589 ATCAGTTAGTCCAGCAATTTCTGTTTCTGTTTGTGTTTGTGATGATGACCTTAATCTTGGAG 648
Db 10112 ATCTGTTAGTTTCACTGTGCTCTGATTAGTTTCTGTTTCCGTGATCTCTGTTGCTG 10053
Qy 649 AGATGGGTATGAGTAGCCCACTATCTGTGTGAGGT-CAATATGTGATTTTACTGT 707

Db 1052 AGAGTGGGGTCTGAAATCTCCCACTATTATTGTATCGGGTGTGATGTGCTTTGAGAT 9993
Qy 708 GTAGCTGTGCTTGTGTTTATGAACCTGGGTGACATGTGTTGGTGGCATAGACATTAAGAA 767
Db 9992 TTAGTAAAGTTTCTTTTATGAATGTGGGTGCCCTTGCATTTGTAGCATACATGTTAGAA 9933
Qy 768 TTGCAATGCTCTCTGGTGGAT-TTCTCTTGGTGGCTATGCTAGTATCTTCCCAATCTC 826
Db 9932 TTGAGAGTTCATCTTGGCAGATGTTTCTTTGACAGATGAAGTGCCTTCTCTTATCTT 9873
Qy 827 ATCTGCTTAGTTTGGGTTT-----AAGTCTATTAGTCAGATATAAAATGACTGTATC 880
Db 9872 TTTTTTTTTGTGATAACTTGGTGGAGTGTGAATTTATTCATATTAGAATGGCTACTCC 9813
Qy 891 GCGTGTGCTCTAGG-----GCCATTTGCTTAGAATA-TCTTTTCCATCTCTTTTACTCTA 934
Db 9812 AGCTTGTCTTGGGAAACAAACCAATTTGCTTGGGAAATTTGTTTCCAACTTGAACCTG 9753
Qy 935 AGGTGATGCTATCCATG---GTAGGTGTCTTTTGGATGCGACAGTAGGATGATCT 991
Db 9752 AGGTAGTCTGCGCTTGTCTACTGAGGTGCATTTCCGTATGCGACAAATGCTGGGTCC 9693
Qy 992 TGTTTTCATATCCATCTGTTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATAGT 1051
Db 9692 TGTTCACACATCCAGTCTGTAGTCTATGTCTCTTTTGGGAAATTTGA-----GTC 9642
Qy 1052 CATTCATGTTGAGAAATATCAATGACGAGTCTTGTGGATCTTGTGTTA-TCTTGCACCTG 1110
Db 9641 CATTCATGTTGAGAAATATTAAGAAAGTGAATTTACTTCTCTGTTATTTTGTGTTG 9582
Qy 1111 TGAAGT 1170
Db 9581 TTAGAGGTGGAATATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9537
Qy 1171 GT 1230
Db 9536 ----- 9537
Qy 1231 CTGGAATATTATTATTATCATATTTTCTTGAATGTGGTAAACATCTTTAGATTGAAGTTT 1290
Db 9536 -TTGAAAGATTACTTCTTCTTCTTCTAGGGTGTAGTTTCCCTCTGTGTGTGTGTGT 9478
Qy 1291 TCTCTCTA--GCTTCTTTAGTCTGTGATTTGAAGATAGATATCTTTTACATCTGATTTT 1348
Db 9477 TCCATCTATTATCTTTTGTAGAGTGGATTTGTGGAATGATATTTGTGTAATTTGGTTT 9418
Qy 1349 ATCTTAGAATGCTTTCTTTCTTCTCAACTATTGTGACAGAAAGTTTCTTAAGTCAGTAG 1408
Db 9417 GTCATGGAATATCTTGGTTTCTTCCATCTATGATTAATTCAGAGTTTGTCTGGGTATAG 9358
Qy 1409 TCTGCGCTGACATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGAGGCGCTTCTTAC 1468
Db 9357 CTTGGGCTGGCATTTGTGTTCTTCTTGGGTCTGTATGACATCTGCTAGGATCTTCTAGC 9298
Qy 1469 ATTTGAGTTCTATTGGAAGAGTCAGGTGAATCTTAATACATCTCCCTTATATCTTA 1528
Db 9297 TTTTATAGTCTCTGGTGAAGAGTCTAGTCTTATCTGTATAGGCTGTGCTTTTATAGTTA 9238
Qy 1529 ATTTGCTTTTTTCCCTTGTGATCTTTTAAATATCTTCTTGTGTTCTATCTATCTTTAGTAT 1588
Db 9237 CTT-GACTTCTTCCCTTACTGCTTTTCAATTTATTTCTTTTGTGTTTAAATGCTTTGATGT 9179
Qy 1589 TTGATATTATGCACTGTGGGAGTTTCTTTTCCGTCCTATCTATTGTTGGTGTGTTGAT 1648
Db 9178 TCAATTATTATGTGACAGAGGA-ATTCTTTTCTGTCCAGTCTATTGAGATTCTGGAG 9120
Qy 1649 GCTTCTGTGCTTGTATAGGATCTCTTCTCAAGTTTAGGAAATTTTCTTTTGTGTT 1708
Db 9119 GTTCTTGTATGTTCAATGACATCGCTTTCTTTAGTTATGGAAGTTTCTTCTATTAAT 9060
Qy 1709 TTTCTGAAAAATTTTCTTCTGCTTTTGTACCT-GCCTTCTCCCTCTCTCTATTC----- 1763
Db 9059 TTGTTGAAGATATTACTTGGCCCTTTGAGTTGGGAATCTTCACTTTCTTCTATACCTATT 9000

Qy 1764 -----TTTGGTTTTTGCATAGTCTCTGGCTTCTGATGCTTTTATGCTCGCTGGATT 1814
Db 8999 ATCCTTAGGTTTCAATTTCTCATTTGTGTCCTATATTTCTGGAATTTTGGTTAGGAGC 8940
Qy 1815 ATTTTAGACTTAACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTCTTCTTCACTG 1874
Db 8939 TTTTTCATTTTGTATTTTCTTTGACTGTGTGTCAATGTTTCTATGATCTTCTGCA 8880
Qy 1875 CCGAGATTTCTCTCTTATCTTCTTGTATCTGTGTCAGTGAGGCTTGTCTGAGTTCTCT 1934
Db 8879 CCGAGATTTCTCTCTTATCTTCTTGTATCTTGTGTTGGTGAATCTGTATATAGTCTCT 8820
Qy 1935 G-TTGGTTTCTTAAATTTTCTTATTTCCAGATTTCTTCTGATTTGGTGTGTTTATTAAT 1993
Db 8819 GATCTCTTCTTAGGTTTCTGCTCCAGCGTTGTCTCCCTTGTGATTTCTTATTTGT 8760
Qy 1994 TCTATTCCACTTTCCAGGTCCTGAATGTTTCTTCTATTTTCTCTCC-----AGTATTACA 2050
Db 8759 TCTAGTTCCATTTTGTAGATCTCTGGATGTTTGTTCATTTCTCTTCCCTGTTTGTATTG 8700
Qy 2051 TTTTTCATAGTTTCTTCTTAAATGATTTTCTTCTCTTCAAGGACCTTTTATGCAAT 2109
Db 8699 TTCTCTGTAGTTCTTTTAAAGATTTTGTGTTTCTCTTTAAGGCTTCTAGTTGTTT 8641

RESULT 34
US-10-087-192-1861/c
; Sequence 1861, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1861
; LENGTH: 175338
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(175338)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1861

Query Match 24.8%; Score 564.9; DB 1; Length 175338;
Best Local Similarity 63.1%; Pred. No. 1.7;
Matches 1159; Conservative 0; Mismatches 551; Indels 127; Gaps 20;
Qy 292 TCTATTCTTGTATCTATCTTGGCTCATTTTAACTAGTAGTGTGTTGTTTCC 351
Db 53933 TTTCTTTCTTTATTTTATATTTGACCAATATATCATTAAGTAGAGCAATGTTTCAGCTTCC 53874
Qy 352 ATAAGTTCTAGTCTTCTGTTTCTGTTGTTGTTTATCTAGATTAAAGCTGT 411
Db 53873 ATGTGATGTGTGTTTCCGTTGTTA---TTGTTGAATTTAAGACCGACCATGCTCTGT 53817
Qy 412 GGTGCTCAGATAGGACATAGAGTATTTTCAATTTGTCTTTTATCTGTGAGACTTGTCT 471
Db 53816 GGTGATCTGATGGGTGTCATGCGATTTTCAATTTCTTTATATCTGTTGAGGCTGATT 53757
Qy 472 TGTTTTGAATATGATTTCAATTTTGGAGA---GTTTCTAGGTTGCTGACAGAGGTAC 529
Db 53756 TGTGACCAATTATAGTGTCAATTTTGGAGAAGGTACCATGAGTGTCTGAAAGAGGTAT 53697

QY 530 AGTCTTTGTTGTTG-GTGAAATAGTCTGTAAATATCT-CTAGGTCCACATTTGGTTTGA 587
Db 53696 ATTCTTTTGGCTTTAGATTGAATGTTCTATAAATATCTGTAGATCCATTTGGTTTCATAG 53637
QY 588 CATCAGTTAGCTCCAGCAATTTCTCTGTTTCGTTTCTTTTGTGAGATGACCTAACTGTTGGA 647
Db 53636 CATCTGTTACTTTCACTGTGCTCTGTTTAG-TTGTGTTTCCATGATCTGTCCATTGCT 53578
QY 648 GAGAAATGGGTATGCAAGTAGCCCACTATCTGTGTGTGAGGT-CAATATGATGATTTAGC 706
Db 53577 GAGAGTAGGGTGTGAAGTCTCCCACTATTTATTTGTGTGTGTGCAAGGTGTGCTTTGAGC 53518
QY 707 TGTAGCTGTCTGTTTATGAACCTTGGGTGACATTTGTTTGTGTGATAGACATTAAGA 766
Db 53517 TTTAGTAAGATTTCTTTATGAATGTGGGTGCCCTTGCATTTGGAGTATAGATGTTTCA 53458
QY 767 ATTCGAATGTCCTCTGTGTGGA-TTTTCTTTGAGTCTATGATATTTCTTCCCAATCT 825
Db 53457 ATTTAGAAATGCACTTGTGTAGATTTTCTTTGATGAGATGAAGTATCTTCTCTTATTC 53398
QY 826 CA---TCTGCTTAGTTTGGGTAAAGTCTATTAGTCAGATATTAAGATGACTGTAT--- 879
Db 53397 TATTTTGTGATACTTTGGTTGAAGTGTGATTTATGATCTAGATGAATGATGACT 53338
QY 880 -CGGCTTCTCTTAGGSCCAATTTGCTTAGAATA-TCTTTTCCATCTCTTTTACTCTAAGG 937
Db 53337 CCAGCTTGTCTTGGGACCAATTTGCTTGGAAATTTGTTTCAAGCCCTTTACTCTGAGG 53278
QY 938 -----TGATGTCATCCATGCTAGGTGCTTTTGGATGAGAGTAGGATGATCT 991
Db 53277 TAGTACTGTCTGTCTTGATCTAGGTGCAATTTCTTTGATGCAACAAATGCTGGGTCC 53218
QY 992 TGTTTTCATATCCATCTCTGTTACCAGTATCTTTTCTAGAGAAATTAAGATCATTTAGT 1051
Db 53217 TACTTACATATCCAGTCTGTTAGTCCATGCTTTTATTAGAATTTAG-----TC 53166
QY 1052 CATTCAGTGTGAGAAATATCAATGAGCAGTGTGTTGAGATCTTGTATCTTGCACTGT 1111
Db 53165 CTTTGTGTTAAGAGATATTAAGGAAAGTGAATTTACTTCTCTATTTATTTTGTGTTA 53106
QY 1112 GAAGT 1171
Db 53105 GAGTGGGAATATGTTGT 53063
QY 1172 TGT 1231
Db 53062 -----T 53062
QY 1232 TGGAAATTTATTTATTCATATTTTCTTGAATGTGGTGAACATCTTTAGATTGAAGTTT 1291
Db 53061 GAAAGAGATATCTTCTGTCTTTTCTAGGTGTAGTTTCCCTCTCTGTGTGTGTGTGT 53002
QY 1292 TCTCTAGCCTTCTT--TAGGTCTGCATTTGAAGATAGATATCTTTACATCTGATTTA 1349
Db 53001 TTGTCTGTATCTTTGTAGGGCTGGATTTGTGAAGAGATATGTGTAAATTTGGTTTG 52942
QY 1350 TCTTAGAATGCTTTCTTTCTCCCACTATGTTGTGACAGAAAGTTTCTTAAGTGCAGPAT 1409
Db 52941 TCATGGAATATTTGTTTCTCTCTATGTTAATTTGAGAGTTTGTGCTGGGTATAGTAC 52882
QY 1410 CTGGCCTGACATCTGTAGTCTTTGGAGTCGTGATGACATCTGTGAGGCCCTCTTACA 1469
Db 52881 CTGAGCTGATATTTGTTTCTCTCAGGCTCTGTATGAGATCTGCCTAAGATCTCTAGCT 52822
QY 1470 TTTTGTAGTTTCTATTTGAAAGAGTCAAGTGTAAATCTAATACATCTGCCTTTATATGTTAA 1529
Db 52821 TTCACAGTCTCT-TGTGAGAAGTCTGTGTTTAAATCTCTGAGAGTCTGCCTTTATATGTTAC 52763
QY 1530 TTGGTCTTTTTCCTCTGCACTTTTAAATATCTTCTTGTGCTATACATTTAGTAT 1589
Db 52762 TT-GACCTTTTCCCTCTTACTGCTTTTAAATATCTTCTTGTGTTTGTGCAATTTGGTGT 52704

QY 1590 TGATTATTATGCACTGTGGGAGTTTCTTTTCCGGTCCAATCTATTTGGTGTGTTTGTATG 1649
Db 52703 TGATTATTATGTGATGGGAGGAATTTCTTTTCTGGTCCAGTCTATTTGGAGTTCTGTAG 52644
QY 1650 CTTCTTGTACCTTGATAGGCACTCTTTTCTCAAGGTTAGGAAATTTTCTTTT---TTG 1705
Db 52643 CTTCTTGTATGTTCAITGGGCACTCTTTTCTTTAGGTTAGGAGTTTCTTCTATACATT 52584
QY 1706 GTTTTCTTGAATAATTTTCCCTGCTTTTGGACCTGCTCTTCCCTTCC-----TCT 1758
Db 52583 TTTTCTTGAAGATAATTTACTGACCTTTAAGTTGTAATTTTCTATACCTTAATATCC 52524
QY 1759 ATTCTTTTGTGTTTTCATAGTCTCTGCTCTTCTGAGTGTGTTTATGCTGCTGATTTT 1818
Db 52523 TTAGATTTGGTCTTCCATTTGCTCTGATTTTCCGGAATTTTCTGAGTTAGAGCTTTT 52464
QY 1819 TAGCAATTAACATTTCTTTGACCAAGATATCCATTTCTTCTATCTTCTTCTTCACTGCTG 1878
Db 52463 TGCATGTTGTGTTTCTTTGACTGTTCTGTGTGATGTTTATATGTTTCTGACCTG 52404
QY 1879 AGATTCCTCTCTCTATCTCTGCTGATCTGCTGAGTGTGCTGCTGAGTTCCTG-TT 1937
Db 52403 AGATTCCTCTCTCTATCTCTGCTGATCTGCTGCGGATGCTTGCATCTAATCTCTGATC 52344
QY 1938 GGTTCTTAAATTTTTCATTTCCAGATTTCTTCACTTTGGGTTTGTGTTTAAATCTTA 1997
Db 52343 TCTTCTTAGGTTTCTAACTCCAGGCTGCTCTCTTTGTATCTTCTTTATTTGTTCTA 52284
QY 1998 TTTCCACTTTGAGTCTGCTGAAATGTTTACTCATTTTCTCT---CCAGTATTTACATTTT 2054
Db 52283 TTTCTCTTTTATAGATCTGAGTGGTTTGTTCATTTCTTCTTCTTCTTATTTGATTTGTTT 52224
QY 2055 CATAGTCTTCTTAAATGAGTTTATTCATTTCTCTCTTCT 2091
Db 52223 CCTGTCTTCTTTAAGGATTTGTGTGTTGTGTTTC 52187

RESULT 35
US-10-322-281-495
; Sequence 495, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 495
; LENGTH: 176594
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(176594)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-495

Query Match 24.5%; Score 558.6; DB 1; Length 176594;
Best Local Similarity 62.5%; Pred. No. 1.8; Mismatches 559; Indels 130; Gaps 19;
Matches 1149; Conservative 0;
QY 292 TCTATTTCTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTAGTGTGTTGTTTCC 351
Db 129540 TTTCTTTCTTTATCTCTTCTTCCCAAGGTATCATTTGAGAAGAGTGTGTTCCGTTTCC 129599
QY 352 ATAAGTTGTAGTTTCTGTTGTTTCTGTTGTTTGTGTTGTTATCTAGATTTAAGCTGT 411
Db 129600 ACGTGAATGTCGCTTTCATTTATCTATGTTGTTTGAAGATC---AGCCTTAGTCCAT 129656
QY 412 GGTGCTCAGATAGGACATAGATTAATTTCAATGTTCTTTTATCTGTCGAGACTTGCTT 471

129657 GGTGATCTGATAGGATCGATGGACAATTTCAATATTTTGTATCTGTTGAGGCGCTGTTT 129716
QY 472 TGTGTTGAAATATGATTCATTTTGGAGA--GTTTCATAGGTCCTGACAGAGGTAC 529
Db 129717 TGTGACCAATATATGTTGCAATTTTGGAGAGGTACCATGAGGTGCTGAGAGAAATGAT 129776
QY 530 AGTCTTTGTTGTTGG--TGAATAGTCTGTAATATCT-CTAGGTCACATTTGTTTATGA 587
Db 129777 ATCCCTTTGTTAGGATGAAATGTTCTGTAATATCTGTTCAATCTATTTGTTTATGA 129836
QY 588 CATCATGTTAGTCCAGCAATTTCTCTGTTGCTGTTTGTGAGATGACCTAACTGTTGGA 647
Db 129837 CTTCCTGTTAGTTTCACTGTGCTGCTGTTTGTAG--TTTCTGTTTCCATGATCTGTTCAITGGT 129895
QY 648 GAGATGGGTATGAGTACCCACATCTGCTGTGAGGT-CAATATGTTGATTTTACG 706
Db 129896 GAGATGGGTGTTGAAATCTCCCACTATTTGTTGAGATGCAATGTTGCTTTGAAC 129955
QY 707 TGTAGTCTGCTGTTTATGAATCTGGGTGACATTTGTTGTTGCTGATAGACATTAAGA 766
Db 129956 TTTACTAAAGTTTCTTTAAATAATGTTGCTGCTGCTTGCATTTGGAGCATAGATATTGCA 130015
QY 767 ATTGCAATGCTCTGTTGATGATTTT-CCTTGATGCTATGATGATTTCTTCCCAATCT 825
Db 130016 ATTGAGATTTCTCTTGAAGATTTTACCTTTGATGATGAAATGTTGCTGCTGCTGCT 130075
QY 826 CATCTGCTTAGTTTGGGTGTTAAGTCTA-TTAGTCAGATATTAATAATGACTGTATCGGCT 884
Db 130076 TTTTGTATTAACCTTTGTTGTTGAGTGGATTTTATCAATATTAGATGGCTACTCCACCT 130135
QY 885 TGTCTTTAGGCGCAATTTGCTTAGAATA-TCTTTTCCATCTTTTACTCTTAAGTGATGT 943
Db 130136 TGTCTTTTCAAGCCATTTGCTTGGAAATTTGTTTCCAGCCTTTCACTCTGAGGTAGTGT 130195
QY 944 CTATCCAT--GGTAGGTTGCTTTTGGATGAGCAGATAGGATGGAATCTGTTTTCAT 1000
Db 130196 CTGCTTTTCCCTGGGATGGGTTTCTGTAAGCAGCAAAATGTTGGTCTGTTGTTGT 130255
QY 1001 ATCCATCTGTTTACCAGTATCTTTTCTAGAGAAATTAAGATCATGAGTCATGATGT 1060
Db 130256 AGCCAGTCTGTTAGTCTATGCTTTTATTTGAGGAAATGAG-----TCCATTTGATAT 130307
QY 1061 TGAGATTTATCATGAGCAGTGTGTTGATTTCTGTTATCTTGCACTTGTGAGTGTGT 1120
Db 130308 TAAGAGATTAAGGAAAGTAATGTTGCTTCTTATTTTGTGTTTAAAGTTGGCA 130367
QY 1121 GTGCTGT 1180
Db 130368 TTCCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 130400
QY 1181 GTGCTGT 1240
Db 130401 -----TTGAAAGAT 130409
QY 1241 TTATTTATTCATATTTTCTGAAATGTTGCTTAACTCTTTAGATTTGAAGTTTCTTCTAGC 1300
Db 130410 TACTTTCTGCTTTTCTAGGTCGTTTCACTTTTATTTGTTTCTTTTCTTTTCT 130469
QY 1301 CTCTTT-----TAGGTCTGATTTGAAGATAGATTTCTTTACATCTCATTTTATCT 1352
Db 130470 GTTGTATCTTTGAAGGCTGATTCATCGAAGATTAATGTTGAATTTGCTTTTCTG 130529
QY 1353 TGAATGCTTTCTTTCTCCCACTATTTGACAGAAAGTTTCTTAAGTCAGTGTCTG 1412
Db 130530 TGAATACCTTTGGTTCTCCATCTAAGTAATTTGAAGTTTGGCTGGGTATAGTACCTG 130589
QY 1413 GCTGACATCTGTAGTCTCTTGGAGTCTGTAGCATCTGTGACGGGCTTTCTTACATTT 1472
Db 130590 GCTGGCATTTGTGTTCTTCTAGTCTGTATAACATCTGTCCAGGCTCTTCTGGCTTC 130649
QY 1473 TGAGTTTCTATTGGAAAGTCAGGTGTAATTTCTATACATCTGCTTTATGTTAATG 1532

Db 130650 ATAGTCTCTGGT-GAAAAAGTCTGGTGTGATTTCTAATAGCCCTGCCTTTATATGTTACTTG 130708
QY 1533 GTCTTTTTCCTTCGATCTTTTAAATATTTCTTTCTTTCTATATCTTTTAGTGAATTTGA 1592
Db 130709 ACCTTTTTCCT--TTATGCTTTTAAATATCTATCTTTATT-TGGTGTGTTTCTGTCGA 130766
QY 1593 TTATATGACACTGTGGGAGTTTCTTTTCCGGTCCAACTATTTGGTGTGTTTGTATGCTT 1652
Db 130767 TTATATGCTGTGAGGAGAAATTTCTTTCTAGTCCAGTCTATTTGGAGTCTGTAGGCTT 130826
QY 1653 CTGTGACCTTTGATAGCATCTTTCTCAAGGTTAGGAAATTTTCTTTTGGTGTGTTTCT 1712
Db 130827 CTGTGATGTTCAAGGCACTCTTTCTTTAGGCAATGGAGTTTCTTCTATATTTTGT 130886
QY 1713 TGAATAATTTTCCCTGCTTTTGAACCTG-----CCTTCTTCCCTTCTCTCT 1758
Db 130887 TGAAGATATTTGCTGGCCCTTTTAAAGTTGAAATCTTTCTTCTGCTACTCTATATAT 130946
QY 1759 ATCTCTTGGTTTTCATAGTCTCTCGCTTCTGATGTTTATGCTGATTTATTT 1818
Db 130947 GTAGGTTTGTCTCTCATTTGCTCTTAAATTTCTTAGAGTTTGACTCAGATCTTTT 131006
QY 1819 TAGACTTAACTTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTGCTGCTG 1876
Db 131007 TGCAITTTGCAITTTTCTTTGATTTGTTGCGCATGTTCTATGTAATCTTCTGCACTG 131066
QY 1879 AGATCTCTCTTCTATCTTGTATTTCTGTCAGTGAAGGTTGCTCTGAGGTTCTCTG 1936
Db 131067 AGATCTCTCTCTTCTATCTTGTATTTGTTGCTAATGCTGCACTATGTTTCTAGAT 131126
QY 1937 TGGGTTCTTAAATTTTCTTCAATTTCCAGATTTCTTCAATTTGGGTTTGTGTTTAAATCT 1996
Db 131127 TTTTCCCTAGGTTTCTATCTTCCAGGTTGCTCATCTTGGGTTTCTTATTTGCTGCT 131186
QY 1997 ATTTCCACTTTGAGGTCCTGAAATGTTTACTATTT--TCTCCAGTATTTTACATTT 2054
Db 131187 ACTTCACTTTTGTAGTCTTGTATGATTTTGTGTAATTTCAATTTCCACCCTGTTGTTGTT 131246
QY 2055 CATAGTCTTCTTAAAGGATTTTCAATTTCTCTCTTCA 2092
Db 131247 TCTGTTTCTTCTAAGGATTTGTTACTGTACTGTACTCTTTA 131284

RESULT 36

US-10-175-523-75
; Sequence 75, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/14795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1

```

; SEQ ID NO 75
; LENGTH: 85548
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-75

```

Query Match 24.4%; Score 556.899; DB 1; Length 85548;
Best Local Similarity 62.1%; Pred. No. 2.5;
Matches 1142; Conservative 0; Mismatches 571; Indels 125; Gaps 18;

QY	292	TCTATTCTTGATTTCTATCTTTGGCTCAATTTTAACTGACGTAGTGTGTTGGTTTCC	351
Db	33322	TTTCTTTCTTTATTCCTCTTGACCAAGGATCATCTGAGAAGAGTGTTTTTCAGTTTCT	33381
QY	352	ATAAGTTTCTAAGTTTTCTGTTGTTTCTGTTGTTGTTGTTTATCTAGATTTAAGCTGT	411
Db	33382	ACATGAATGTTGGCTTTCTATTATTTATTTCTATTGAAGATC---AGCCTTAGTCCAT	33438
QY	412	GGTGGTCAGATAGGACATAGAGTATTTATTTCAATTGCTCTTTTATCTGTCTCGAGACTGTCT	471
Db	33439	GGTGATCTCATAGGATGCATGGGCAAAATTTCAATATTTTGTGATATGTGAGGCTGTGTT	33498
QY	472	TGTTTTGAAATATGATTTCAAATTTTGGGAGA--GTTTTATAGGGTCTCTCACAGAAGGTAC	529
Db	33499	TGTGACCAATATATGCTGCAGTTTTTGGAGAAGGTACCATGAGTGCTCGAGAAGAGGCAT	33558
QY	530	AGTCT-TTTGTGTTTTGGTGAAATAGTCTGTAATAATCT-CTAGTGCCACTTGGTTTTATGA	587
Db	33559	ATCCCTTTGTTTTAGGATAAAATGTTCTGTAGATATCTCTAGATGCCATTTGTTTCATAA	33618
QY	588	CATCAGTTAGTCCAGCATTTCTCTGTTTCTGTTTTTGTGTGAGATGACCTAACTGTTGGA	647
Db	33619	CCTCTGTAGTTTCACTGTGTCCTCTGTTTAG-TTTCGTGTTTCCATGATCTGTCCANTGGT	33677
QY	648	GAGAATGGGGATATGAAGPAGCCCACTATCTGTGTGTGAGGTCAATATGTCAATTTTAGCT	707
Db	33678	GAAAGTGTGTGTGGAAGTCTCCCACTATTATTGTGTGAG-----GTGCTTTGAGCT	33729
QY	708	GTAGCTGTCTTGTTTTTATGAACITTTGGGTGACATTTGCTTTGGTGATAGACATTAGAA	767
Db	33730	TTACTAAGTTTCTTTTAAATGAATGGCTGTCCTTCTATTGGAGCATAGATGTCAGAA	33789
QY	768	TTGCAATGTCCTTTGGTGGATTTT-CCTTTGANGCCTATGTAGTATTCTTTCCCAATGTC	826
Db	33790	TTGATAGTTCTCTTTGGAGATTTTACCTTTGATGAGTATGAAGTCTCTCTCTGTGCTT	33849
QY	827	ATCTGTGTTATTTGGGTTTAGTCTA-TTAGTCAGATATTAAATGACATGTATCGGCTT	885
Db	33850	TTTTGATGACTTTGGTTGGGAAGTCGATTTTATTCGATATTAGAATGGCTATTCCATCTT	33909
QY	886	GCTTCTTAGGGCCATTGCTTTAGAATA-TCTTTTCCATCCCTTTTACTCTAAGGTGATGTC	944
Db	33910	GTTTCTTCTATATCATTTGCTTGAAAAATGTTTTTCCAGCCCTTTCATCTCGAGGTAGTGC	33969
QY	945	TATCCAT---GGTAGGTTGCTTTTTTGGATGTCAGCAGTAGATGGATCTTTGTTTTTCATA	1001
Db	33970	TATCTTTTCTCTGAGATAGATTTCCCTGTAAAGCAGCAAAATTTGTGGGTCGTGTTGTGTA	34029
QY	1002	TCCATTCTGTTACCCAGTATCTTTTTCTAGAGAAATTAAGATCATGTAGTCATTGATGTT	1061
Db	34030	GCCAGTCTGTAGTCTATTTCTTTTT-----TATTGGGAGTTGAGTCCATTGATTT	34082
QY	1062	GAGAAATATCAATGACGAGTGTTTGTGGATCTCTGTTATCTTGTGCACTTGTGGAAGTGTG	1121
Db	34083	AAGAGATATTAGGAAAAAGTAAATCTTGCTTCCATTAT-----	34121
QY	1122	TG	1181
Db	34122	-----TTTTGTTGAAGTTGGCATTCGTCTTGTGTGTGTGTGTGTGTGTGTGTGTG	34167
QY	1182	TGTCGTGTGTGTGTGTGTGTGTGTCTCTCCCTCTTTTGTATTTTGTGGCCTGGAAATATT	1241
Db	34168	GGTTGTTGTGGGGAATACCTTTTGTCTTTTCTTAGGACATGTGTTTCTGTGCTTGTATTGTT	34227

1242	TAATATTCATATTTCTTGGAATGTGGGTAAACAATTTAGATGAAGTTTTTCTCCTAGCC	1301
34228	TTTTTTT-----TTCTTGTATTATC	34248
1302	TTCTTTAGGTCGTGATTTGAAGATAGATATCTTTTACATCTGATTTTATCTTGAAGATGTC	1361
34249	CTTTGAGGGCTGGATTCATGGAAGAATAATGTGGAATTTGGTTTGTGTGGAAATACT	34308
1362	TTTCTTTCTCCAACTATTGTGACAGAAAGTTTTTCTAAAGTCAGAGTAGTCTGGCCCTGACAT	1421
34309	TTGGTTTCTCCACTATGAGTAATTCAGAAATTTGGCTGGGTATAATATTTTGGGTGGCGT	34368
1422	CTGTAGTCCTCTGGAGTCTGTAGCACATCTGGCAGGCGCTCTTACATTTTGGATTCT	1481
34369	TTTGTTTCTCTTAGTGCTTTTATAACATCTGTCCAGGCTGTCTGGCTTTCATACTCT--	34426
1482	ATTGAAAAAGTCAGGTGTAATTTCTAATACATCTGCCTTTATATGTAAATGGTCTTTTTT	1541
34427	-GGTGA AAAAATCTGGGTAAATTCGTAGGCTTGCCTTATATGTACTT-GCCCTTTTA	34484
1542	CCCTTGCATCTTTTAAATATCTTCTTTTGTCTATACHTTTTAGTGAATTGATTAATATGC	1601
34485	CCCTTGTGCTTTTAAATATCTATGTTTGTGTAGTGCAATTTGTTGTTCTGAATTAATGT	34544
1602	ACTGTGGGAGATTTCTTTTCGGTGCCAATCTAATTTGGTTGTTTGTATGCTTCTTGTACCT	1661
34545	GTCCGGAGGAATTCCTTTCTGGTCCAGCTCTATTGGAGTTCGTAGGCTTCTTATATGT	34604
1662	TGATAGGCATCTCTTTCTCAAGGTTAGAAATTTTCTTTTTTGGTTTCTGTGAAAATAT	1721
34605	TGATGCCATCTCTTTCTTTAGGTTTGGGAAGTTTTCTTCTATATTAATTTGTTGAAGATAT	34664
1722	TTTCCCTGCTTTTGACCTG-----CCTTCTCCCTTCTCCTATATCCCTTTG	1767
34665	TTGCTGGTCTTTGAGTTGAAAAATCTTCATCTCATCTACTCTATTATCCGTAGGTTTG	34724
1768	GTTTTTGATAGTGTCTCTGGCTTCTCGATGTTTTTATGCTCGGATTAATTTAGACTTAA	1827
34725	GTCTTCTCATTTGTGCTCGGATTTCTGGATGTTTGAGTTAGGATCTTTTGGCATTTTC	34784
1828	CATTTTCTTTGACCAAGGATCCATTTCTTCTATCTGTCTTCACTGCCCTGAGATTTCTCT	1887
34785	CCTTTTCTTGATGTTGTGTGATGTTCTCTATGGAATCTTTCTGACCTGAGATTTCTCC	34844
1888	CTTCTATCTCTTGATTTCTGCAGTAGGCTGTGCTCTGAGGTTCCCTG-TTGGGTTCTTA	1946
34845	CTTCCATCTCTTGATTTCTGTGTGTGGTCTCGCTTCTGTGGTTCCAGATTTCTTTCCCTA	34904
1947	ATTTTTTCATTTCCAGATTTCTTCAGTTTGGGTTTTTGTTTATTAATTCATTTTCCACTT	2006
34905	GGGTTTCTACCTCCAGCGTTGCCACATTTGGGTTTTCTTTATTGTGTCTACTTCCCTTT	34964
2007	TCAGGTCTCGAAATGTTTTTACTCATTTTCTCCCTCC--AGTATTTTACATTTTTCATAGTTT	2063
34965	TTAGGCTTGAATGGTTTTTATTCAAATCCATCACTCTTTGGTGTGTTTTTCTCGCAATT	35024
2064	CTTTTAAATGGATTTATTCATTTTCTCTTCAAGGACCTTT	2101
35025	CTTTAAGGGATTTTGTCTTCTCTTTTAAAGGTTCTCT	35062

RESULT 37

RESULTS 37
US-10-052-482-133

US 10 052 402 133
; Sequence 133, Application US/10052482
; Publication No. US20040072264A1

GENERAL INFORMATION:

; APPLICANT: Engelhard, Eric

APPLICANT: MORRIS, David

; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER

```

; FILE REFERENCE: A-71087/RMS/DCF

```

; CURRENT APPLICATION NUMBER: US/10/052,482

; CURRENT FILING DATE: 2002-08-15

Db 31891 -----TCTCTGTATCTGTGGTGATGCTTGCATCTAAGACTCCTGATCTCTTCCC 31942
QY 1945 TAAATTTTTCATTTCCAGATTTCCCTCAGTTGGTGTGTTTGTATTAATTAATTTCCAC 2004
Db 31943 TAGGTTTCTATCCGGTGTGCTCCCTTGGATTTCTTTATTTGTTCTATTTCCA- 32001
QY 2005 TTTCAGTCTCGAATGTTTACTACATTTTCTCC---AGTATTTACATTTTCATAGGT 2061
Db 32002 TTTTAGATCTTGGATAGTTTGTTCATTTCTCTCCCTGTTTGTATGTTGTTTCTCTGTAA 32061
QY 2062 TTTCTTAATGATTTTATTCATTTCTCTCTCAAGGACCTTTTATGAATTCATAA 2114
Db 32062 TTTCTTAAGGAGATATTATGCTCTTCTTAAAGTCTCTATCATCATCA 32114

RESULT 38
US-10-087-192-1765
; Sequence 1765, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1765
; LENGTH: 68370
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(68370)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1765

Query Match 24.2%; Score 550.599; DB 1; Length 68370;
Best Local Similarity 62.0%; Pred. No. 2.9;
Matches 1153; Conservative 0; Mismatches 584; Indels 124; Gaps 19;
QY 292 TCTATTTCTGATTTCTATCTGCTCAATTTTAACTCAGTAGTGAAGTTGTTGTTCC 351
Db 45984 TTTCTTTCTTTATTCATCTTGCACCAAGGTATCATTTGAGAAGAGTGTATTTCTGTTTCC 46043
QY 352 ATAGCTTTGTAGTTTCTGTTGTTTCTGTTGTTGTTGTTGTTTATCTAGATTTAAGCTGT 411
Db 46044 ATGTTAAAGTTGGCTTTTCCAAATATTATGTTGTCAATTTGAAGATG---AGCCTTAGTCCAT 46100
QY 412 GGTGCTCAGATAGGACATAGATATTATTTCAATTTGCTTTTATCTGTCGAGACTTGTCTT 471
Db 46101 GGTGCTCTGATAGGATGATGGACAAATTTCAATATTTTGTATCTGTTGAGGCTGTTT 46160
QY 472 TGTTTGAATATGATTTCAATTTTGGAGA--GTTTCAAGGGTGGTGCACAGAAGGTAC 529
Db 46161 TATGACCAATATATATGTTCAATTTTGGAGAAGGTCCCGTAGAGTGTCTGAGAAGAAGGTAT 46220
QY 530 AGTC-TTTGTGTTTGGTGAATAGTCTGTAATAA-TCTCTAGTCCCACTGGTTTATGA 587
Db 46221 ATCTTTTGTGTTTAGGATATAATGTTCTGTAGATACCGGTTAAGTCAATTTGTTTCATAA 46280
QY 588 CATCAGTTAGTCTCAGCAATTTCTGTTTCTGTTTCTGTTTGTGAGATGACCTAATCTGTTGA 647
Db 46281 CTCTGTTAGTTTCACTGTGCTCTGTTTAG-TTTCGTGTTTCCAGACTCTGCTCATTTGAT 46339
QY 648 GAGATGGGTATTTGAAGT-AGGCCACTATCTGTGTGAGGT-CAATATGATTTTAG 705

Db 46340 GAAAGTGTGTGTGTAATTTCTCCCACTATTATTGTGTGAGGTGCAATGTGTGCTTTGAG 46399
QY 706 CTCTAGCTGTGCTTGTGTTTATGAACTTGGTGACATTTGTTGGTGACATAGACATTAAG 765
Db 46400 CTTTACTAAGGTCTCTTAAATGATGTGCTGCCCTTGCATTTGGGCAATAGATATTCA 46459
QY 766 AATTGCAATGCTCTCTGTGTGGAATTTT-CCTTTGATGCCATGATGATATCTTCTCCAAATC 824
Db 46460 AATTGAGAGTCTCTCTGGAGGATTTTACCTTTGATGAGTATGAAGTGTCCCTCTCTGTG 46519
QY 825 TCATCTGCTTGTGTTTGGGTTTAAAGCTA-TTAGTGCAGATATTAAATAACGACTGTATCGGC 883
Db 46520 TTTTGTGATAACTTTGGGTTGGAATGATTTTTCATATACGAATGGCTACTCCAGC 46579
QY 884 TTGCTTTCTTAGGGCCATTTGCTTAGAATA-TCTTTTCCATCTCTTTTACTCTAAGGTGATG 942
Db 46580 TTGTTTCTTCAGACCAATTTGCTTTGGAATAATTTTCCAGCCCTTTCACCTCTGAGGTAGTG 46639
QY 943 TCTATCCAT---GGTAGGTGTCTTTTGGATGACGAGTAGGATGATCTGTTTTCATCA 999
Db 46640 TCTGTCTTTTCCCTCGAGATGGGTTTCCCTGTAAAGCAGCAAGATGTTGGGTCTGTTGTG 46699
QY 1000 TATCCATTTCTGTACCAGTATCTTTTCTTAGAGAAATTAAGATCATTTGAGTCATTTGATG 1059
Db 46700 TAGCCATTTCTGTAGTCTATGCTTTTATTTGGGAATTTAG-----TCCATTTGATA 46751
QY 1060 TTGAGAAATTAATCAATGAGCAGGTTTGTGGAATCTTTGTTATCTTGCATCTTGTGAAGTGTG 1119
Db 46752 TTAAGAGATATTAAAGCAAAAGTAATTTGCTTCCCTTTTATTTT----- 46795
QY 1120 TG 1179
Db 46796 ----- 46795
QY 1180 TGTGTCTGTCTGTGTGTGTGTGTGTCTCTCCCTCTTTTGTGATTTTGGCCTGGAATTA 1239
Db 46796 --TGTTGTTATAGTTGGCATTTCTTTCTGTGGCAGTCTCTTTTGTGTTTATTTAGGGGA 46853
QY 1240 TTTATTTATTCATATTTTCTTGAATGTTGGTGAACATCTTTAGATGGAAGTTTCTCTCTAG 1299
Db 46854 TTACTTTCTGTTGTGTTCTAGGGCGGTATTCGTCCTTTGATTTGTTCTTTCTGTTAT 46913
QY 1300 CCTTCTTT--AGGCTGCAATTTGAAGATAGATATTCTTTACATCTGATTTTATCTTAGAA 1357
Db 46914 TATCCTTTGAAGGCTGGATTTCTTGGAAAGATTAATGTTGAAATTTGTTTGTCTGTTGAA 46973
QY 1358 TGTCTTTCTTTCTCCAACTATTGTGAGCAAGATTTTCTTAAGTGCAGTAGTCTGCGCTG 1417
Db 46974 TACTTTGCTTTCTCCATCTATGTTAATTTGAGAGTTTGGCTGGGTATAGTAGCTGGGCTG 47033
QY 1418 ACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGACAGGCTTCTTTACATTTTGTAGT 1477
Db 47034 GCATTTGTGAATCTCTTAGTGTCTGTATAACATCTGTTCAGGCTCTTCTGGCTTTCTAGT 47093
QY 1478 TTTCTATTGAAAAAGTCAGGTGAATTTCTAATACATCTGCTTTTATATGTTAATTTGTTCT 1537
Db 47094 CTCTGCT-GAAAAAGTCTAGTGAATTTCTGATAGGCTGCTTTATATGTTACTT-GACCT 47151
QY 1538 TTTTCCCTTGCATCTTTTAAATTTCTTTTGTGTTCTATACCTTTAGTCAATTTGATTTAT 1597
Db 47152 TTTCTCCCTTACTGCTTTTAAATTTCTATCATTAATTTAATGCAATTTGTTTCTGATTTAT 47211
QY 1598 ATGCACTGTGGGAGTCTTTTCCGTCCTCAATCTATTTTGGTGTGTTTGTATGCTTCTGT 1657
Db 47212 ATGTGCTCGAGGAATTTCTTTTCTGCTCAGTCTATTTGGAGTCTGTAGGCTCTTCT 47271
QY 1658 ACCTTGATAGGCAATCTTTCTCAAGGTAGGAATTTTCTTTTGGTGTGTTTCTGTTGAA 1717
Db 47272 ATGATCATGGGCACTCTTTTCTTTATGTTGGGAAGTTTCTTCTATTTATTTGTTGAG 47331
QY 1718 ATATTTTCCCTGCTTTTGGACCTG-----CCTTCTTCCCTTCTCTATTTCC 1763
Db 47332 ATATTTTCTGGCCCTTTAAGTTGAAATCTTCAATCTCAATCACTCTTATTAATCCGTAGG 47391

Db 43096 TCCCTTGAAGGCGTGAATTCGTGGAAGATAATGTGAAATTTGGTTTGTGTCATGGAATA 43155
QY 1360 TCTTTCTTCTCCAACTATGTGACAGAAAGTTTCTTAAGTCAGATAGTCTGGCCGTGAC 1419
Db 43156 CTTTGGTTTCTCATCTATGTAATGGAAG-TTTGCTGGGTATAGTACCTATGCTGGC 43214
QY 1420 ATCTGAGTCTCTTGAGTCTGTAGACATCTGTGACAGGCGCTTCTTACATTTTGAAGTTT 1479
Db 43215 ATTCTGTCTCTTAGTGTCTGTATATAACATCTGTCCAGGATCTTCAGACTTTCATAGTCT 43274
QY 1480 CTATTGGAAGTCAAGTGTAAATCTTAATACATCTGCTCTTATATGTAATTTGTCCTTTT 1539
Db 43275 CTGGT-GAAAAGTCTGGTGTAAATCTGTATAGGCTTGGCTTTATTTGTTACTTGACCTTTT 43333
QY 1540 TTCCCTTGATCTTTTAATATTTCTTCTTCTTCTTCTATATCTTTTGTGATGATTTGATTTAT 1599
Db 43334 TTCCCTTACTGCTTTAATATTTCTATCTTTATTAGTGTACTTGTGTTCTGATTTAT 43393
QY 1600 GCACGTGGGAGTTCTTTCCGGTCCAACTATTGCTGTGTTTGTATGCTTCTTGTAC 1659
Db 43394 GTGTCAAGGAGAAATTTCTTCTGCTCCAGTCTATTGGAAGTCTGTAGGCTTCTTGTAT 43453
QY 1660 CTGTATAGCATCTCTTCTCAAGGTAGGAAATTTTCTTTTGTGTTTCTTGAATAAT 1719
Db 43454 GTTCATGGGCATCTCTTCTTCTGTTTGGGAAGTTTCTTCTATATTTTGTGAAGAT 43513
QY 1720 ATTTTCCCTGCTTTTGAAGCTG-----CCTTCTTCCCTTCTCTATTCCTT 1765
Db 43514 ATTTCTGCGCCCTTTAAGTTGAAAATCTTCAATCTCATCCACTCTTATTCATAGCTT 43573
QY 1766 TGGTTTTTGCATAGTCTCTGCGCTTCTGATGTTTATGCTGGATTTATGCTGGATTTTGTAGCTT 1825
Db 43574 TGGTCTTCTCATGTGTTCTGGAATTTCTGATGTTTGTAGTTAGGATCTTTTGTGATTT 43633
QY 1826 AACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCACTGCTCGATTTCT 1885
Db 43634 TGCAATTTTCTTTGATTTGTGTCATGTTCCCTACAGAACTTCTGCACTGACATTTCT 43693
QY 1886 CTCCTTCTATCTCTGTATCTGTGATGAGGCTGTCTGAGGTTCTCTG-TTGGGTTCT 1944
Db 43694 CTCCTTCCATCTCTGTATCTTGTGTTGCTGATCTGATGTTCTGATTTCTTCTTCTTCC 43753
QY 1945 TAAATTTTTCATTTCCAGATTTCTTCAATTTCTTCTATCTTGTCTTCACTGCTCGATTTCT 2004
Db 43754 TAGGTTTCTATCTCCAGTGTGCTCACTTGGTTTCTTATTTGTGCTACTTCCCT 43813
QY 2005 TTTCAAGGCTGAAATGTTTACTCATTTTCTTCCAGTATTTCTTCCAGTATTT---TACATTTCTATAGT 2061
Db 43814 TTTGAGGCTAGTATAGTATTTTGTTCATTTCCATCACCTGTTTGGATTTGCTTCTCTGTTT 43873
QY 2062 TCTTTTAATGATTTATTTTCAATTTCTTCTTCAAGGACCTTTTATGAAATTCATPAAATGAT 2121
Db 43874 TCTTTAAGGACTTCTACTGTTGATTTGTTGTTTCTTCTGTTTCTTTAAGGACTTGTAA 43933
QY 2122 GTTAAG 2127
Db 43934 GTCAAG 43939

RESULT 41
US-10-175-523-62/c
; Sequence 62, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/J7795-US3

; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 185548
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-62

Query Match 24.0%; Score 546,799; DB 1; Length 185548;
Best Local Similarity 62.2%; Pred. No. 1.9;
Matches 1156; Conservative 0; Mismatches 572; Indels 132; Gaps 20;

QY 292 TCTATTTCTTGATTTCTATCTTTGGCTCAATTTTAACTCAGTAGTGAGTTGTTGGTTTCC 351
Db 129820 TTTCTTTCTTTATTTCTTCTTGAACCAAGTGTCAATTAAGAAGAGTGTGTTTCAGTTTCC 129761
QY 352 ATAAGTTCTGAGTTCTTCTGTTTCTGTTCTGTTGTTGTTTATCTAGATTTAAGCTGT 411
Db 129760 ACTGAATGTGGCTTTCATTTATTTATG-----TGTTATTGAAGATCACCCCTAGTCCAT 129705
QY 412 GGTGGTCAGATAGGACATAGAGTATTTATTTCAATTTGCTTTTATCTGTCGAGACTTGCTT 471
Db 129704 GGTGGTCTGATAGGATGATGGCAATTTCAATTTTGTATCTGTAGGCTGTGTT 129645
QY 472 TGTTTTGAATATGATTTCAATTTTGGAGA--GTTTCTAGGGTCTGCACAGAGGTAC 529
Db 129644 TGTGACCAATATATGTTCAACTTTGGAGAAGTCCCATGAGTGTCTGAGAAGAAGGTAT 129585
QY 530 AGTC-TTTGTTGTTTGGTGAATAGTCTGTAATATCT-CTAGGTCCACTTGGTTTATGA 587
Db 129584 ATCCCTTTGTTTGAATATAATGTTCTGTAGATATCTGTTAAGTCCATTTCTTTCAATA 129525
QY 588 CATCAGTTAGTCCAGCATTTCTGTTTCTGTTTCTGTTTGTGAGATGACCTAACTGTTGA 647
Db 129524 CTTCTGTTAGTTCATCTGTGCTCCCTGTTT-GGTTTCTGTTTCCATGATCTGTCCATTGAT 129466
QY 648 GAGAATGGGTATGGAAGTAGCCCACTATCTGTGTGTGAGT-CAATATGATTTTATAGC 706
Db 129465 GAAAGTGTGTGTGAAGTCTCCCACTATATTTGTGTGAGTGTCAATGTGTGCTTTGAGC 129406
QY 707 TGTAGCTGTGCTTGTGTTTATGAACCTTGGGTGACATGTTGTTGGTGCATPAGACATTT-AAG 765
Db 129405 TTTACTAATGTCTTTTAAATGAATGTGGTGCCTTGCATTTGGAGCATAGTATTTTCAGG 129346
QY 766 AATTGCAATGCTCTCTTGGTGGATTTT-CCTTTGATGCTATGATGATTTCTTCCCAATC 824
Db 129345 AATTGAGAGTTCTCTCTTGAAGGATTTTACCTTTTATGAATGAAAGTGTCCCTCTCTGTC 129286
QY 825 TCATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATATAAATGACTGTATCGGC 883
Db 129285 TTTTGTGATACTTTTGGGTGGAAGTCGATTTTATTCGATATTAGAATGGCTACTCCAGC 129226
QY 884 TTGCTTCTTAGGGCCATTTGCTTAGAATA-TCTTTTCCATCTCTTTTACTCTAAGGTGATG 942
Db 129225 TTGTTTCTTTCAGACCGTTTGTCTGGAAATTTGTTTCCAGCCCTTTTCACTCTGAGTAGTG 129166
QY 943 TCT--ATCCATGGTAGTGTCTTTTGGATGTCAGCAGTAGGATCTTGTGTTTCTAT 1000
Db 129165 TCTTCTTTTCCCTGAGATGGGTTTCCCTGTAAGCAGCAGAAATGTTGGTTTCTGTTTGTGT 129106

QY 1001 ATCCATTCTGTACCCGATCTTTCTAGAGAAATTAAGATCATTTAGTCAATGATGT 1060
Db 129105 AACGAGTCTGTAGTCTAAGCTTTTATTTGAGAAATGAG-----TCCATGATAT 129054
QY 1061 TGAGAAATATCAATGAGCAGTGTGTGAGATCTTTGATCTTTGACATCTGGAAGTGTGT 1120
Db 129053 TAAGAGATATTAAAGAAAAGTAATGTGTCTTCTTTTATTTGTTAGAGTGGCA 128994
QY 1121 GT 1180
Db 128993 TTTCTTTCTGTGGCTGTCTTCTTTTGGTTG----- 128961
QY 1181 GT 1240
Db 128960 -----TTGAGGGAT 128952
QY 1241 TTAATATTCATATTTCTTGAATGAGGTGAACATCTTTAGATTGAAGTTTCTCCTAGC 1300
Db 128951 TACTTTCTTACTTGTCTAGGCGGTGATTTCTGCTTGTATGCTTCTTTCTGTATT 128892
QY 1301 CTTCTTT--AGTCTGCAATTTGAAGATAGATATTTTACATCTGATTTATCTTAGAAT 1358
Db 128891 ATCCTTTGAAGGGCTGATTTGTTGAAGATATTTGTGAAATTTGGTTTCTGCGAAT 128832
QY 1359 GTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAAGTCAGTAGTCTGCCCTGA 1418
Db 128831 ACTTTGGTTTCTCCATCTATGTAATTTGAGAGTTTGGCCGGTATAGTAGCTGGCTGG 128772
QY 1419 CATCTGAGTCTCTGAGTCTGTAGCACATCTGTGAGGCCCTTCTTACATTTGATTT 1478
Db 128771 CATTTGTCTCTTAGTGTCTGTATAACATCTGTGACAGGCTCTTCTGGCTTTCATAGTC 128712
QY 1479 TCTATTGAAAGTCAGGTGTAATCTTAATACATCTGCTTTATATGTTAATTTGGTCTTT 1538
Db 128711 TCTGCT--GAAAGTCTGTGTAATTTCTCATAGGCTTCTTTATATGTAATTT--GACCTT 128654
QY 1539 TTTCCCTTGCATCTTTTAATATCTTTCTTTGTTCTATACATTTTGTAGTGAATGATTA 1598
Db 128653 TCTCCCTTACGCTTTTAATATCTTATCTTTTATTTAGTGAATTTGTTGTGATTA 128594
QY 1599 TGCAGTGTGGGAGTTCTTTTCCGGTCCAATCTATTGTTGGTTTGTATGCTTCTTTGTA 1658
Db 128593 TGTGTGGGAGAAATTTCTTTTCTGGTCCAGTCTATTGGAGTTCTGTAGGCTTCTTGTGA 128534
QY 1659 CTTGTATAGGAATCTTTTCTCAAGGTTAGAAATTTTCTTTTCTTTTCTTTTCTTTGAAA 1718
Db 128533 TGATCATAGGCATCTCTTTTATGTTTGGAAAGTTTCTTCTATTATTTTGTGAAGA 128474
QY 1719 TATTTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCTCTATTCTCT 1764
Db 128473 TATTAGCTGGGCCCTTTAAGTTGAATACTTCAATCTCATCAATCTCTATTATCCGTAGGT 128414
QY 1765 TTGGTTTTGATAGTGTCTCTGGCTTCCCTGGATGTTTTATGCTCGGATTTATTAGACT 1824
Db 128413 TTGGTCTTCTCATTTGTCTGATTAACCTGA-----GTTAGGATCCTTTTGCAAT 128362
QY 1825 TAAATTTTCTTTGACCAAGTATCCATTTCTTCTATCTTCTTCTTCTTCTTCTTCTTCTTCT 1884
Db 128361 TTGATTTTCTTTGACTGTGTGTGATGTTCTCTATGGAATCTTCTGCACTGAGATTC 128302
QY 1885 TCTCTTCTATCTTTGATTTCTGTGAGGCTGTCTCTGAGTTTCTCTG--TTGGGTTCT 1943
Db 128301 TCTCTTCAATTTCTGTGATTTCTGTGTGATGCTGATGCTGATGTTTCCAGATCTCTTTC 128242
QY 1944 TTAATTTTCTCAATTTCCAGATTTCTTCAAGTTTGGTTTTGTTTATTAATTTCTATTCCCA 2003
Db 128241 CTAGGGTTTCTATCTCCAGCGTTGCCCTTGGGTTTTCTTTATTTGTGCTGCTTCCC 128182
QY 2004 CTTTCAGTCTGAAATGTTTTTACTCATTTTCTCCAGTATT---TACATTTTCTAGG 2060
Db 128181 CTTTATGTTCTAGTATGTTTTTGTTCATTTCCATCCACTGTTTGGATGTGTTTCTCTGTT 128122

QY 2061 TTTCTTTAATGAGTTTATTATTCATTTCTCTCAAGGACCTTTTATGATTAATAAATGTA 2120
Db 128121 TTTCTTTAAGATTTCTACTGTTGGTGTGTTTCTCTGCTTTTCTTTAAGGCGCTGTA 128062
RESULT 42
US-10-087-192-235
; Sequence 235, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 235
; LENGTH: 91071
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(91071)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-235

Query Match 23.6%; Score 537,898; DB 1; Length 91071;
Best Local Similarity 63.0%; Pred. No. 2,9;
Matches 1180; Conservative 0; Mismatches 551; Indels 143; Gaps 24;
QY 292 TCTATTTCTGATTTCTATCTTTGGCTCATTTTAACTCAGTAGTGTGTTGGTTTCC 351
Db 24602 TTTCTTTGTTTATTTCTTCCCTGACCAAGTTGTCTATGTAGTAGAGAGTTGTTCC 24660
QY 352 ATAAAGTTTGAAGTTTCTGTTTCTGTTGTTGTTGTTATCTAGATTTAAGCTGT 411
Db 24661 ACTTTTATGTGGGATTTATGTTATTTTCTGTTTATTTGAAGA--CCAGCCTTAGGCCAT 24717
QY 412 GTGTGTCAGATAGGACATAGATTTATTTCAATTTGCTTTTATCTGTGAGACTTGTCT 471
Db 24718 AGTGGTCTGATAGATACATGTTATTTATTTTCTGTTATCTGTTGAGGTTGTTT 24777
QY 472 TGTTTTGAATATGATTTCAATTTTGGAG--AGTTTCATAGGTCCTGACAGAAGGT-A 528
Db 24778 TTTGACTAATTATATGTTTGGTCAATTTTGGAGACAGTACCATGAGGTGCTAAGAAAAAGATCT 24837
QY 529 CAGTCTTTCTGTTTGGTGAATAGTCTCTAAATATCT--CTAGGTCCTACTTGGTTTATCA 587
Db 24838 CTTCCCTTTGTTTAAAGTAAATGTTCTATAGATATCTGTTAAATCCATTGATTCATAA 24897
QY 588 CATCAGTTAGTCCAGCATTTCTGTTTCTGTTTCTGTTTGTGTGAGATGACCTAACTGTGA 647
Db 24898 CCTCTATTAGTTTCACTGTTTCTGTTTCTGTTTGTAG--TTTCTGTTTCAATGACCTGTCCATGGGT 24956
QY 648 GAGATGGGATTTGAGTAGCCCATCTCTGTTGTGAGG--TCAATATGATGATTTTAC 706
Db 24957 GAAATGGAGATTTGATGCTCTGACTATATTGTTGGGGTCAATGTTGTTGTTTAC 25016
QY 707 TGTAGCTGTGTTTGTATTAAGAACTTGGGTGACATTTGTTTGGTGCATAGACATTAAGA 766
Db 25017 TTTATTAAGCTTCTTTTATAAATCTGTGATGCCCTTGTATTTTGGGCAATAGATTTAGA 25076
QY 767 ATGCAATGCTCTTGTGGGATTTTCTTGTG---ATGCTATGTAGTATTTCTCCCAAT 823
Db 25077 ATTGAGACTTCTCTGTGGTGGATTTTCTCCTATGATATATATGTAATGACCTTCTCAT 25136

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/374,077
 FILING DATE: 25-Feb-2003
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Roseman, Stephen
 REGISTRATION NUMBER: 43,058
 REFERENCE/DOCKET NUMBER: 100107.401D1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 209:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 51259 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 SEQUENCE DESCRIPTION: SEQ ID NO: 209:
 US-10-374-077-209

Query Match 22.9%; Score 521.797; DB 1; Length 51259;
 Best Local Similarity 60.7%; Pred. No. 4.3;
 Matches 1118; Conservative 0; Mismatches 637; Indels 88; Gaps 17;

QY	292	TCTATTCTGCTGCTATCTTGGCTCATTTTAACTCAGTAGTGAGTGTGTTGGTTCC	351
DB	26099	TCTTCATTTTCATTTCTGCCCTGACCAAGTATCATTCAGTAGAGCGCTGTTCAGCTTC	26158
QY	352	ATAAGTTTGAAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT	411
DB	26159	ATATGATGTGTGCTTCCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	26218
QY	412	GGTGTGATAGGACATAGAGTATTTTCAATGTCTTTTACTCTGTCGAGACTTCGTT	471
DB	26219	GGTGTGATAGGAGTGTGAGTATTTTCAATGTCTTTTACTCTGTCGAGACTTCGTT	26278
QY	472	TGTTTGAATATGATTTCAATTTTGGAGA--GTTTCATAGGGTGTGTCAGCAAGAGTAC	529
DB	26279	TGTGACAGCTATATGTCAGTTTGGAGAGGTTTCCATGAGGTGTGAGAGAGGTAT	26338
QY	530	AGCTTTGTTGTTTGGTGAATAGTCTGTAATATCT--CTAGTTCACCTGGTTTATGAC	588
DB	26339	ATTTTTCGTTTGGATGACATGTTCTATAAATATCTGTTAGATCCCATTTGGTTCAATAC	26398
QY	589	ATCAGTTAGTCCAGCATTTCTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT	648
DB	26399	ATCTGTTAGTTTCACTGTCTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTT	26458
QY	649	AGAAATGGGATATGAGTAGGCCATATCTGTGTGTGAGGTGCAATATGTGATTTAGCTG	708
DB	26459	AGAGTGGGCTGCTGAAATCTCCCATATATTTGATCATCAGGTATGATGTGCTTTGAGAT	26518
QY	709	TAGCTGTCTGTTTATGAACTGGGTGACATTTGTTGTTGTTGTTGTTGTTGTTGTTGTT	768
DB	26519	TTAGTAAAGTTTTTTTATGAAATGGGTGCCCTTGCAATTTGGAGCATACATGTTTCAGAT	26578
QY	769	TGCAATGTCCTCTGTTGGAT--TTTCTTTGATGCTATGATGATTTCTTCCCAATCTCA	827
DB	26579	TGAGAGTTTCATCTTGGCAGATGTTTCTTCTGACCAATATGAGTGTCTCTCTTATCTTT	26638
QY	828	TCTGCTTAGTTTGGTTTAACTCT--ATTAGTCAGATATTAATGACTGTATCGGCTTG	886
DB	26639	TTTTTGAATCTGTTGAGAGTTGAAATTTATTTCCATATTAAGATGGTACTCCAGCTTG	26698
QY	887	CTTCTTAGG--GCCATTTCTTTAGAAATA--TCTTTTCCATCTCTTTTACTCTAAGGTGA	940
DB	26699	TTTCTTGGGAACAACCAATTTGCTTTGGAAATTTGTTTCCCACTTTGAATCTGAGGTAG	26758
QY	941	TGCTATCCATG--GTAGTTGCTTTTGGATGAGCAGTAGGATGGATCTTGTGTTT	997
DB	26759	TGCTGCTCTTGTCTACTGAGGTGCAATTTCTGATATGACGCAAAATGCTGGGTCTGTGTTA	26818

QY	998	CATATCCATTCTGTTTACCCAGCATCTTTTCTAGAGAAATTAAGATCATTCAGTCAITGA	1057
DB	26819	CACACCAGTCTGTTAGTCTATGCTTTTGTGAGAAATGA-----GTCCATTTGA	26869
QY	1058	TGTTGAGAAATATCAATGAGCAGTCTTTTGTGGATCTTTGTTATCTTGCACCTTGCAAGTG	1117
DB	26870	TGTTAAGAGATATTAAGGAAAGTCACTTACTTCTCTGTTATTTT-----TGTTG	26920
QY	1118	TGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1177
DB	26921	TTGTTAGAGGTGAATATATGTTTGTGGCTATCTCTTTTGGTTGTTGTTGTTGTTGTTGTTG	26980
QY	1178	TCTGTTCTGTTCTTGTGTTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1237
DB	26981	TTTCTGCTTTTCTAGGGGTAGTTTCCCTC-----CTTGTGTTGTTGTTGTTGTTGTTGTTG	27025
QY	1238	TATTTATTTATCATATTTTCTTGAATGTTGGTAAACATCTTTAGATTGAATTTTCTCTCT	1297
DB	27026	TTTCCATCTATTTATCTTTTGTAGAGCTGGAAGATATTTGTGTTGTTGTTGTTGTTGTTGTTG	27085
QY	1298	AGCCTCTTTTGTGTTCTGCAATTTTGAAGATAGATATTTCTTACATCTGATTTTATCTTAGAA	1357
DB	27086	AAATACCTTAGCAGCTTGACAGCACACCTCAACACCTAGACCTAAAGAGCAATACAC	27145
QY	1358	TGTTTCTCTCTCTCAAACTATTTGTGACAGAAAGTTTCTTAAGTCGAGTAGTCTGGCTG	1417
DB	27146	CCAAGAGGAGTAGACTGAGATTGGGAGTT-----TTGCTCTGGGCTG	27186
QY	1418	ACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGSCCTCTTACATTTTGTAGT	1477
DB	27187	GCATTTGTTCTCTTGTAGGCTGTGTGACATCTGCTAGGATCTTTTGTAGCTTTTGTAGT	27246
QY	1478	TTCTATTTGAAAAGTCAGGTGTAATTTCTAATACATCTGCTCTTATATGTTAATGCTCTT	1537
DB	27247	TTCTGTG--GAGAAGCTGGTGTAAATTCATGATAGGCTGCTCTTATATGTTACTTGACCT-	27304
QY	1538	TTTTCCCTTGCACTTTTATATTTCTTTTGTCTATCTTTTGTAGTATTTGATTTT	1597
DB	27305	TTTCCATTTGCTCTTTAATATTTCTTTTGTGTAGTCAITTTGGTGTGTTTGTATTT	27363
QY	1598	ATGCACTGTGGGAGTTCTTTTCCGGTCCAACTATTTTGTGTTGTTTGTATGCTCTTGT	1657
DB	27364	ATGTGACAGAGGAATTTCTTTCTGCTCAGTCTATTTGGAGTCTTGAGGCTCTTGTC	27423
QY	1658	ACCTTGATAGGAATCTTTTCTCAAGTTAGAAATTTTCTTTTGTGTTTCTTCTGAA	1717
DB	27424	ATGTTCACTGGCATCGCTTTTGTAGTTAGGAAAGTTTCTTCTCTATATTTTGTGAG	27483
QY	1718	ATATTTTCCCTGCTTTTGTAGCTGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1763
DB	27484	ATATTTTCTGCTTTGAGTTGGGAACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	27543
QY	1764	TTTGGTTTTGATAGTGTCTCTGCTCTCTGATGTTTATGCTGCTGATATTTTAGAC	1823
DB	27544	TTTGGTCTTCTCATTTGCTGCTGGAATTTCTGCTGATTTTGGGTTAGGAGCTTTTGTGAT	27603
QY	1824	TTAACATTTTCTTGACCAAGTATCCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1883
DB	27604	TTTGTATTTTCTTGTACTGTTGTGTCATATTTCTATGTTGATCTCTGACCTGAGATT	27663
QY	1884	CT	1941
DB	27664	CT	27723
QY	1942	TCTTAATTTTCTTCT	2001
DB	27724	TCCTAGATTTCTTAACCTCCAGGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	27783
QY	2002	CACCTTCAGCTCTGAAATGTTTTTACTCATTTTTTCTCTCCAGTATTTTACA--TTTTCATA	2058
DB	27784	CATTTTAGACCTGGAATGTTTTTGTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	27843

Qy	2059	G G T T C T T T A A T G G A T T T A T T C A T T T C C T C T T C A A G G A C C T T T	2101
Dδ	27844	T A A T T C T G T A A G A A T T T T T G T G T T T C C T C T T T A A G G G C T T C T	27886

RESULT 45

```

US-10-085-117-61/c
; Sequence 61, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
;   APPLICANT: Morris, David W.
;   APPLICANT: Engelhard, Eric K.
;   TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
;   FILE REFERENCE: 529452000121
;   CURRENT APPLICATION NUMBER: US/10/085,117
;   CURRENT FILING DATE: 2002-02-27
;   PRIOR APPLICATION NUMBER: US 09/798,586
;   PRIOR FILING DATE: 2001-03-02
;   NUMBER OF SEQ ID NOS: 361
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 61
;   LENGTH: 96602
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;     NAME/KEY: variation
;     LOCATION: (1)..(96602)
;     OTHER INFORMATION: n = any nucleotide
US-10-085-117-61

```

Query Match	22.9%	Score 521.599	DB 1	Length 96602
Best Local Similarity	61.8%	Pred. No. 3.3		
Matches 1157	Conservative 0	Mismatches 579	Indels 136	Gaps 22
Qy	289	ACTCTATTCTTCGATTCTCTATCTGGCTCAATTTTAACTTAACTCAGTAGTGAGTTGTTGGTT	348	
Db	5698	AAITTCCTTCTTATTTCTTCCCTTGACCAAGTATCATTCAGTAGAGTGTGTTCACT	5639	
Qy	349	TCCATAGTTTGTGAAGTTTTCTGTGTGTTCTGTGTTGTTGTTGTTATCTAGATTTAAGC	408	
Db	5638	TCCATGTATGTGTGCTTCCATTTGTTTGGTATTAAAG---CCATCCTTAGTC	5582	
Qy	409	TGTGGTGGTCAGATPAGACATAGAGTATTATTTCAATTGTCTTTTATCTCTCGAGACTTG	468	
Db	5581	AGTGGTGGTGTGATAGGTCGATGGGGTTATTTCAATCTCTTATATATGTTGAGACCTG	5522	
Qy	469	CTTTGTTTTGAAATATGTTATTTCAATTTTGAGA--GTTTCATAGGGTGCTGACAGAAGG	526	
Db	5521	TTTTTGACCAATTATATGTCATTTTGAGAAGGTACAGTAGGTGCTTGAGAAGG	5462	
Qy	527	TACAGTCTTTGCTGTTTTGG--TGAAATAGTCTGTAATAATC-TCTAGGTCACCTTGTTTTA	584	
Db	5461	TATATTTCTTTCCCTTAGGATGAAATCTTCTATAAATATCAATTAGATCTATTTTGGTTCA	5402	
Qy	585	TGACATCAGTTAGTCCAGCATTTCTCTGTTTCTGTTTTGTTGTTGAGATGACCTAACTGTT	644	
Db	5401	TAACTTCTGTTAGTTTCACTGTCCTGTTTAG-TTTCTGTTTCCATGATTGTCCATT	5343	
Qy	645	GGAGAGTAATGGGTATTGAATAGTCCCACTATCTGTGTGTGAGGTCAATATGTGATTTTA	704	
Db	5342	GCTAAGAGTGGAGTGTAAAGTCTCCCACTATTATTAGTGGGGGT-GCCATGTGTGCTAA	5284	
Qy	705	GCTGTAGTCTGCTGTTTATGAACTTGGGTGACATGTGTGTTTGGTGCATAGACATTAA	764	
Db	5283	GCATTAGTAAAGTTTCTCTTATGAATGTGTTGCCCCTTGCATTTTGGAGCATAGATGTTCA	5224	
Qy	765	GAATTCGAATGCTCTTGTGTGA-TTTTTCTTTGTAGCCATGTAGTATTTCTTCCCAAT	823	
Db	5223	GAATTGAGATTTATTTTTGTGTAGATTTTTCTTTGACCAGTATAAAGTGTCCCTCCATT	5164	
Qy	824	CTCATCTGCTTAGTTTGGTTTTAA--GTCTATTAGTCAGATATTAATAATGACTGTATCG	881	
Db	5163	CATTTTTTGATAACTTTTGGTGTGAAGTGTGATTTTTATTAGATATTAGAATGACTACTCCA	5104	

QY	882	GCTTGTCTCTTAGGGCCATTGCTTAAGATA-TCTTTTCCATCTTTTAACTCTAAAGGTGA	940
DB	5103	GCTTGTCTCTTAGGGACCGTTGCTTGAAATTTGTTTCCAACTCTTTACTCTCGAGGTAG	5044
QY	941	TGTCATACCCATG---GTAGGTGCTCTTTTGTGATGCAGCAGTAGGATGCATCTCTTTT	997
DB	5043	TGTCGTCTCTTAGGCACCTGAGATGCACCTCTGTATGCAGCAAAATGCGGGTCTCTTTA	4984
QY	998	CATATCCATCTCTGTATCCCAAGTATCTTTTCTAGAGAAATTAAGATCAITTAGT-CATTG	1056
DB	4983	TGTATCCAGTCGGTTATTCTATGTC-TTT-----TTTGATAATTGAGTCCATTG	4935
QY	1057	ATGTTGAGAAATATCAATGAGCAGTGTGTTGTGGAATCTGTGTAATCTTGCACTGTGAAGT	1116
DB	4934	ATGTTTAGAGATATTAAGGAAAAGTAATGTTACTTCCCTGTTATTTTGTGTAGAGGC	4875
QY	1117	GTGTGCTGT	1176
DB	4874	AGAATTATGTTGTGTGGCTATCTTCTTTTGGGTTG-----	4838
QY	1177	GTCTGTGCTGT	1236
DB	4837	-----TTGAA	4833
QY	1237	TTATTTATTTATCATATTTTCTTGAATGUGGTAACATCTTTAGATTGAAGTTTTTCTCC	1296
DB	4832	AGATTACTTCTTGTCTTTTCTAGGGTGTAATTTCCCTCTGTGTGTGTGTGTGTGTGTGT	4773
QY	1297	TAGCCTTCTT--TAGGTCTGCATTGAAGATAGATATCTTTACATCTGATTTTATCTTA	1354
DB	4772	TATTATCCTTTGTAGGGCTGGAATTAAGAAAGATATGTGTAAA-TTAGTTTGTTCATG	4713
QY	1355	GAATGCTTTTCTTCTCCAACTATTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTGGC	1414
DB	4712	GAATATCTTGTTTTCTCCACTAATGTAATGAGATTTGTGCGGTGATATAAGACATGGG	4653
QY	1415	CTGACATCTGTAGTCTCTTGGAGTCTGTGACATCTGTGCAGGGCCTCTTACATTTTG	1474
DB	4652	CTAGCATTTGTGTTCTCTTAGGGTCTGTATGACATC-ATCCAAAATCTCTAGCTTCAT	4594
QY	1475	AGTTTCTATTGGAAAAGTCAGGTGTAATCTTAATACATCTGCCTTTATATGTTAATGGT	1534
DB	4593	AGACTCTGGT-GAGAAAGTCTGGGTAGTTCTGATAGGTTTGGCTTCTATATGTACTGAC	4535
QY	1535	CTTTTTTCCCCTTGATCTTTTAAATATCTTCTTCTTCTTCTATACATTTTAGTCATTGAT	1594
DB	4534	CTTTTTT-CCTTACTGCTTTTAAATATCTTCTTCTTGTGTAGTCATTGGTGTTTTGAT	4476
QY	1595	ATTATGACCTGTGGG---GAGTTCTTTTCCGGTCCAACTATATTGGTGTTTGTATGC	1650
DB	4475	ATTATGAGACAAGGGGAATGAATTTCTTTTCTGCTCCAGTCTATTGTGAGATTCTGTAGGC	4416
QY	1651	TTCTTGTACCTTCATAGGATCTCTTCTCAAGGTTAGGAATTTTCTTTTGGTTTT	1710
DB	4415	TTCTGTATGTTTCATAGGATCTCTTCTTTAGGTTAGGAATTTTCTCTATAAATTT	4356
QY	1711	CTTGAATAATTTTCCCTGCTTTTGCCTG-----CCTTCTCCCTCTCCT	1756
DB	4355	GTTGAAGATATTTCTGGCCCTTTTGAGTTTGAAATCTTTTGTCTCTCTATACCTATTAT	4296
QY	1757	CTATTCCTTTGGTTTTTGATAGTAGTGTCTCTGGCTTCTGGATGTTTTATGCCGTGAATAT	1816
DB	4295	CCTTAGGATTCATCTCTCAITGTGTCTCGAGATTTTCTCGATATTTTGGGTTAGGAGCT	4235
QY	1817	TTTAGACTTAACATTTTCTTGACCAAGGTATCCATTTCTTATCTCTCTCTCACTGCC	1876
DB	4235	TTTGTATTTGANTTTTGTGACTGTGTGTCAATGTTTTCTATGGTATCTCTGTGCACC	4176
QY	1877	TGAGATCTCTCTCTCTATCTCTGTATCTCTGACGTAGGCGTGTCTGTAGGTTCTGTG-	1935
DB	4175	TGAGATCTCTCTCTCTATCTTTTGAAGTCTGTGTGTGATGCTTTGATCTATGACTCCTGA	4116

us10664775-4.rnpb

Wed Aug 25 15:11:59 2004

Qy	1936	TTGGGTTCTTAATTTTTCATTTCCAGATTCCTTCAGTTTGGGTTTGTATTATTAATTC	1995
Db	4115	CTCTTTCTAGGTTTCTAACTCTACGGTTGTCTCCCTTTG---TACTTTCATTAATTC	4059
Qy	1996	TATTTCCACTTTCAGGTCCTGAAATGTTTACTCATTTTCTCC---AGTATTTACAT	2052
Db	4058	TATTTCCATTTTGAATCTTGGATGTTTGTTCATTTCCATCACCTGTTGATGTT	3999
Qy	2053	TTGATAGGTTTCTTTAATGGAATTTATTCATTTCTCTTCAAGGACCTTTTATGAATTCAT	2112
Db	3998	TTCTCTAAATCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT	3939
Qy	2113	AAATGTATGTT	2124
Db	3938	TTGTTTGTGTTT	3927

Search completed: August 25, 2004, 09:28:53
Job time : 4466 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 25, 2004, 09:30:42 ; Search time 151 Seconds
(without alignments)
2.623 Million cell updates/sec

Title: US-10-664-775-4

Perfect score: 2279

Sequence: 1 gatacctcctctagtgaag.....ttgtaattctagtgctgat 2279

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 45 seqs, 86898 residues

Total number of hits satisfying chosen parameters: 90

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rst4.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	558.599	24.5	3843	1 AK041981	ACCESSION:AK041981
c	2 556.099	24.4	3887	1 AK087425	ACCESSION:AK087425
	3 550.401	24.2	2841	1 AK046480	ACCESSION:AK046480
	4 547.599	24.0	3235	1 AK033883	ACCESSION:AK033883
	5 544.798	23.9	3025	1 AK037785	ACCESSION:AK037785
	6 544.399	23.9	4895	1 AK085507	ACCESSION:AK085507
c	7 531.598	23.3	3204	1 AK085877	ACCESSION:AK085877
	8 521.598	22.9	1644	1 AK089113	ACCESSION:AK089113
	9 516.198	22.7	2436	1 AK089845	ACCESSION:AK089845
	10 515.998	22.6	4832	1 AK045811	ACCESSION:AK045811
	11 503.498	22.1	3375	1 AK045866	ACCESSION:AK045866
c	12 481.397	21.1	4466	1 AK034168	ACCESSION:AK034168
	13 457.898	20.1	2108	1 AK041956	ACCESSION:AK041956
	14 448.597	19.7	3179	1 AK048889	ACCESSION:AK048889
	15 430.098	18.9	4502	1 AK079106	ACCESSION:AK079106
	16 428.398	18.8	2278	1 BC021486	ACCESSION:BC021486
c	17 418.399	18.4	3279	1 AK040345	ACCESSION:AK040345
	18 412.599	18.1	2535	1 AK050942	ACCESSION:AK050942
	19 399.299	17.5	1151	1 B2124878	ACCESSION:B2124878
	20 393.099	17.2	1065	1 B2135754	ACCESSION:B2135754
	21 385.199	16.9	2372	1 AK040990	ACCESSION:AK040990
c	22 363.399	15.9	886	1 B2109340	ACCESSION:B2109340
	23 361.399	15.9	846	1 B21348678	ACCESSION:B21348678
	24 358.399	15.7	825	1 B2221947	ACCESSION:B2221947
	25 355.899	15.6	900	1 B2132018	ACCESSION:B2132018
	26 354.899	15.6	1037	1 B2108508	ACCESSION:B2108508
c	27 353.899	15.5	914	1 B2135754	ACCESSION:B2135754
	28 351.899	15.4	1075	1 B213244	ACCESSION:B213244
	29 351.499	15.4	920	1 B2108872	ACCESSION:B2108872
	30 351.099	15.4	823	1 B2109050	ACCESSION:B2109050
	31 349.899	15.4	816	1 B2184947	ACCESSION:B2184947
c	32 349.699	15.3	910	1 B2097017	ACCESSION:B2097017
	33 349.499	15.3	1061	1 B2244766	ACCESSION:B2244766

34 349.099 15.3 830 1 B2272430
c 35 348.299 15.3 878 1 B2103057
c 36 347.299 15.2 2559 1 AK088420
c 37 346.499 15.2 835 1 BZ097284
c 38 346.299 15.2 833 1 BZ106482
c 39 346.299 15.2 843 1 BZ099109
c 40 345.299 15.2 854 1 BZ109162
c 41 345.099 15.1 838 1 BZ091343
c 42 345.099 15.1 849 1 BZ258120
c 43 344.999 15.1 821 1 BH327408
c 44 344.899 15.1 831 1 BH258418
c 45 343.899 15.1 762 1 BH357910

ALIGNMENTS

RESULT 1
AK041981
LOCUS
DEFINITION
AK041981 3843 bp mRNA linear HTC 19-SEP-2003
Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched
library, clone:A630050F23 product:Rho-associated coiled-coil
forming kinase 1, full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK041981.1 GI:263334838
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
AUTHORS
TITLE
JOURNAL
MEDLINE
MEDLINE
PUBMED
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
REFERENCE
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
5
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

6
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
7
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3843)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiranoto, K., Hiraoka, T., Hirczane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Db 2300 TAGGTTGGGCGAGTTTCTTCTAATAATTTGTTGAAGATATTTACTGGCCTTTAATTT 2359

Qy 1740 G-----CCTTCTTCCCTTCTCTATCTTCTTGGTTTGTGATAGTGTCTC 1785

Db 2360 GGAAACCTTCCTCTCGTCTATACCTATATCTTAGTGTGCTTCTCTATTGTGCT 2419

Qy 1786 TGGCTTCTCGATGTTTATGCTTGCCTGGATATTTTAGACTTAACATTTTCTTACCAAG 1845

Db 2420 GGATTTCTCGATGTTTGTGTTAGGATCTTTTTCATTTTTCATTTTCTATCTGATGTTG 2479

Qy 1846 TATCAATTTCTTATCTTGTCTTCACTGCTGAGATCTCTCTCTATCTTCTTGTATTC 1905

Db 2480 TGTCAATGTTTCTATGTTATCTTCTGCACTGAGATCTCTCTTCCATCTCTTGGATTC 2539

Qy 1906 TGTCAAGGAGCTTGTCTCTGAGATTCCTG-TTGGGTTCTTAAATTTTCTTATTCAGAT 1964

Db 2540 TGTGGTATGCTTGCACCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2599

Qy 1965 TTTCTTCAAGTTGGTGTGTTTATTAATCTTATTCCTTCCACTTTCAGCTCTGAAATGTTT 2024

Db 2600 TGGTCTCCCTTGTGATGTTCTTCTTATTTTCTTCTGTTTCTTCTTCTTCTTCTTCTTCT 2659

Qy 2025 TACTCAATTTCTCTCCAGTATTTTACATTTTCTATAGGTTTCTTAAATGATTTATCA 2081

Db 2660 TGTTCATTTCTTTCACCTGATGATGTTCTCTCTGTTATTTCTTAAAGGATTTATTA 2719

Qy 2082 TTTCTCTTCTCAAGACCTTTTATGATATCAATAATGATGTTTAAAGTCTCTGCTTGTG 2141

Db 2720 TTTCTCTTCTTAAAGTCTTCTATCATCATCATGAGAGTGAATTTAGATCCATATCTTGT 2779

Qy 2142 CTTCAGCTATGTTGATCTCTCAGGCTT 2169

Db 2780 TTTCTGATGATGATATATCAGACT 2807

RESULT 2

AK087425

LOCUS

DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone: E130020L12 product: unknown EST, full insert sequence.

ACCESSION AK087425

VERSION AK087425.1 GI:26104279

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL 20499374

MEDLINE 11042159

REFERENCE 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitasuna, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL

20530913

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6

AUTHORS (bases 1 to 3887)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Iotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nomura, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/

FEATURES

source

1. 3887

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:E130020L12"

/db_xref="MGI:2425267"

/db_xref="taxon:10090"

/clone="E130020L12"

/tissue_type="eyeball"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="0 day neonate"

misc_feature

1. 3887

/note="unknown EST (GB|BB797679, evidence: BLASTN, 96%, match=243)"

Query Match 24.4%; Score 556.099; DB 1; Length 3887;

Best Local Similarity 62.0%; Pred. No. 0.7;

Matches 1144; Conservative 0; Mismatches 574; Indels 127; Gaps 18;

Qy 292 TCTATTCTGATTTCTATCTTGGCTCAATTTAACTCAGTAGTGTGTTGGTTCC 351

Db 720 TTTCTTTCTTTATTTCTTCTTCCCTTCCAGCAGGATCATTTGAGAGAGTGTTCAGTTCC 779

Qy 352 ATAAGTTTCTAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 411

Db 780 ACGTGAATGTGGCTTCTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 836

Qy 412 GGTGGTCAATAGGACATAGATGATTTATTTCAATTTCTTTTCTTTCTGTCGAGATTTGCT 471

Db 837 GGTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 896

QY	472	TGTTTGGAAATATGATTTCAATTTTGGAGA--GTTCATAGGGTGCTGACAAGAAGGTAC	529
DB			
DB	897	TGTGACCAATTATATGGTCAGTTTGGAGAAAGGTACCATAGGGTGCTCGAGAAGAAGGTAT	956
QY	530	AGTC-TTTGTGTTTTGGTGAAATAGTCTGTAAATATCT-CTAGGTCACATTTGTTTATGA	587
DB	957	ATCCTTTTGTTTTAGGATAAAGTGTTCTGTAGATACTGTGTAATCCAGTGTGTTCTATAA	1016
QY	588	CATCAGTTAGTCCAGCAATTTCTGTGTTTCGTTTTTTTGTGTAGATGACCTAACTGTTTGA	647
DB	1017	CTTCTGTTTGTTCACGTGTCCTCGCTTAG-TTTCTGTTTCTATGATCTGTCCATTGAT	1075
QY	648	GAGAAATGGGGTATGAAGTAGCCCACTATCTGTGTGAGGT-CAATATGTGATTTTACG	706
DB	1076	AAAAGTAGTTTATGTAATCTCCACCATATTATTATGTAGAGGTCAATGTGTGTTTGAGC	1135
QY	707	TGTAGTGTCTCTTGTTTTATGAACTTGGGTGACATTTGTTTGGTGATAGACATTAAGA	766
DB	1136	TTTAAATAAAGTCTCTTTAATAAATGTGGTCTTGTGTGCAATTTGGAGCATAGATTTCCAG	1195
QY	767	ATTGCAATGTCCTCTGGTGGAATTTCC-TTTGATGCTATGTAGTATTCTTCCCAATCT	825
DB	1196	ATTGTGAGTGCTCTTTGGAAAGTTTACTTTTGATGAGTATGAAGTGTCCCTCCTGTCT	1255
QY	826	CATCTGCTTGTAGTTTGGGTTTAAAGTCTA-TTAGTCAGATATTAAAAAGACTGTATCGGCT	884
DB	1256	TTTTTGATTAACTTTGGAGTGGAACTCAATTTTATTGATATTAGAATGACTCTCCAGCT	1315
QY	885	TGCTTCTTAGGGCAATTGCTTAGAATA-TCTTTTCCATCCTTTTACTCTAAGGTGATGT	943
DB	1316	TGTTTCTCAGACCAATTGCTTGGAAATTTGTTTCCAGCTTTTCACTCTGAGGTAGTAT	1375
QY	944	CTATCCAT---GGTAGGTGTCTTTTTTGGATGCAGCAGTAGGATGGATCTTGTTTTCAT	1000
DB	1376	CTCTATTTTCCCTGAGTGGGTTTCTGTGAGCAGCAAAATGTTGGTGCTCTGATGTGT	1435
QY	1001	ATCCATCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTTGAGTCAATGATGT	1060
DB	1436	AGCCAGTCTGTTAGTCTATGCTTTTTTATTGGGGAATTTAG-----TCCATTTGATAT	1487
QY	1061	TGAGAAATTCAATGACAGTGTTTGGGAATCTTGTATCTTGCACTCTGTGAAGTGTGT	1120
DB	1488	CAGAGACATTATAAAGAATAATTTTGCTTCTTTTATTGTTTGTAGAGTTGGCA	1547
QY	1121	GT	1180
DB	1548	TTCTGTCTTTGTGGCTGTCTTCTTTTAGTGTG-----	1580
QY	1181	GTGTGTGTGTGTGTGTGTGTCTCTCCCTCTTTTGATTTTTTGGCCTGGAATTAT	1240
DB	1581	-----TTAAGGAT	1589
QY	1241	TTATATTATCATTTTCTTGAATGGGTAACTTTTAGAATGAAGTTTTTCTCTAGC	1300
DB	1590	TACCTTCTTCTTTCTAGGATGGTTTCCAACTTGTATTGTTTTTTTTTTTTTTTCT	1649
QY	1301	CTTCTTTTAGTCTGCAATTTCAAGATAGAT----ATTCTTTACATCTGATTTTATCTTAG	1355
DB	1650	GTTATTATCTTTGAAGGCTGGAATTCATGGAAAGATGTGTGAATATGGTTTGGTGTGG	1709
QY	1356	AATGTCTTCTTCTTCCAACTATGTGACAGAAAGTTTTCCTAAGTGCATGTCTGGCC	1415
DB	1710	AATGCTTTGGTTTCTCCATCTATGTAATTGAAAGTTTGGCTGGGTATAGTACCAGGGC	1769
QY	1416	TGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGACAGGCCCTTCTTACATTTGA	1475
DB	1770	TGGCATTTGTGTCTCTTAGTGTCTGTATAACATCTGTCCAGGATCTTCTGGCTTTCATA	1829
QY	1476	GTTTCTATTGGAAGAAGTCAGGTGTAATCTTAATACATCTGCCCTTATATGTTAATTTGGTC	1535
DB	1830	GTCTCTGGT--GAAAAGTCTGGTGAATTCATGTAGGCCCTGCCCTTATATGTTACTT--GAC	1887

Qy	1536	TTTTTTCCCTTGACATCTTTTAATATCTTTTCTTTGTTCTATACCTTTTACGTGATTTGAATTA	1595
Db	1888	CCCTTTCCCTTACTGCTTTTAAATATCTCTCTTTATTTAGTAGCATTTGTTGTTCTGAATTA	1947
Qy	1596	TTATGCACCTGTGGGAGATTCCTTTTCCCGGTCCAAATCTATTTGSGTGTTTGTATGCTTCTT	1655
Db	1948	TTATGTGTGGGAGGAATTCCTTTTCCCGGTCCAGTCTATTTGGAGTTCGTAGGCTTCTT	2007
Qy	1656	GTACCTTGATAGGCACTCTTTTCTCAAGGTAGGAAATTTTCTCTTTTGGTTTTCTTGA	1715
Db	2008	GTATGTTTCATGGCATCTCTTTCTTTAGGTTCAGACATTTTCTCTATAAATTTTGTTGA	2067
Qy	1716	AAATATTTCCCTGCTTTTGTGACCTGCTCTCTTCCCTTC-----CTCTAT	1760
Db	2068	AGATATTTGCTGCTCTTTAAAGTTGAAAACCTCTTCATCTCTCATCTACTCTATATCTGT	2127
Qy	1761	TCTTTTGGTTTTGCAATAGTCTCTGGCTTCCTGGATGTTTTATGCCGTGGATATTTTA	1820
Db	2128	AGGTTTGGTCTTCTCATGTGCTCGAATTCCTCGATGTTGAGTTGGGATCTTTTGTG	2187
Qy	1821	GACTTAACATTTCTTTTGACCAAGGTATCAATTTCTTCTATCTTGCTTCTCACTGCCTGAG	1880
Db	2188	CATTTTGCAGTTCTTTTGATTTGTGTGGTGAATTTCTCTATGGAACTCTTCGCACCTGAG	2247
Qy	1881	ATTCTCTCTCTATACCTTGTATTTCTGTGAGTAGGCTTGCTCTGAGGTTCTGT--TTGG	1939
Db	2248	ATTCTCTCTCCATCTCTGTATTTCTGTGTCTGATGCTCGCATCTATGGTTCCAGATTTTC	2307
Qy	1940	GTTCTTAAATTTTTCATTTCCAGATTTCTTTCAGTTTGGGTTTGTGTTTAAATTCATTT	1999
Db	2308	TTTCTTAGGGTTTCTATCTCCAGCGTTGGCTCACTTTGGCTTTTCTTATTTGTGCTACT	2367
Qy	2000	TCCACTTTCAGGTCCTGAAATGTTTTTACTCATTTTCTCCC--AGTATTTACATTTTCA	2056
Db	2368	TCCTTTTTAGGCTTTGGATGGTTTTATTCAAATTCATCACCTGTTTGGTCGTGTTTCC	2427
Qy	2057	TAGGTTTTCTTTAATGGAATTAATTCATTTCTCTTCAAGGACCTTT	2101
Db	2428	TGCAATCTTTTAAAGGATTTATGCTGTCTCTTTTAAAGTCTCTCT	2472

RESULT 3	AK046480/c	LOCUS	DEFINITION
----------	------------	-------	------------

AK046480 2841 bp mRNA linear HTC 20-SEP-2003
Mus musculus adult male corpora quadrigemina cDNA, R1EN
full-length enriched library, clone:523039724 product:similar to
L1 RETROPOSON, ORF2 MENA (FRAGMENT) [Rattus norvegicus], full
insert sequence.

insert sequence.
ACCESSION AK046480
VERSION AK046480.1 GI:26091544
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE	1	Manamata, Suchanara, Kocencova, Zdenek, Carninci, P. and Hayashizaki, Y.
AUTHORS		High-efficiency full-length cDNA cloning
TITLE		Meth. Enzymol. 303, 19-44 (1999)
JOURNAL		99279253
MEDLINE		10349636
PUBLISHED		PUBMED

- AK046480 2841 bp mRNA linear HTC 20-SEP-2003
Mus musculus adult male corpora quadrigemina cDNA, RIKEN
full-length enriched library, clone:B20397K24 product:similar to
L1 RETROPOSON, ORF2 MRNA (FRAGMENT) [Rattus norvegicus], full
insert sequence.
- AK046480 AK046480.1 GI:26091544
HTC; CAP trapper.
- Mus musculus (house mouse)
- Mus musculus
- Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Sciurognathi; Muridae; Murinae; Mus.
- 1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253 10349636
- 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374 11042159
- 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Teshi, Y., Nakamura, S., Hayada, M., Nishino, T., Hayada, A.,


```

QY 1409 TCTGGCCTGACATCTGTAGTCTCTTGGAGTCGTAGACATCTGTGCGAGCCCTCTTAC 1468
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1214 CTTGGGTTGGCATTTGTCTCCCTTAGGCTGCGAATCATCTGTCCAGGATCTTCTGGC 1155
QY 1469 ATTCTGAGTTCTATTGGAAAGTCAGGTGTAATCTTAATCATCTGCTTTATATGTTA 1528
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1154 TTTCATAGTCTCTGTG- GAAAAATCTGGTGTATCTGATAGGCTGCTTTATATGTTA 1096
QY 1529 ATTGTCCTTTTCCCTTGCATCTTTTAATATCTTTCTTTGTTCTATPACTTTTAGTGAT 1588
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1095 CTT-GACCTGTTTCCCTTACTCTCTTTAATATCTTATCTTTATTTAGTGCAATTTTGTGTT 1037
QY 1589 TTGATTAATATGACACTGTGGGAGTTCTTTTCCGTCCTCAATCTATTTTGGTGTGTTGAT 1648
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1036 CTGATTAATATGTCGGGAGGAATTTCTTTTCTTGTCCAGTCTATTTGGAGTCTCTGTAG 977
QY 1649 GCTTCTTGTACCTTCATAGGATCTCTTTCTCAAGGTTAGGAATTTTCTTTTGTGTT 1708
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 976 GCTTCTTGTATGTCATGSSCATGTCATTTTAGGTTTGGAGTTTCTTCTATATTT 917
QY 1709 TTTCTGAAAAATTTTCCCTGCTTTTGACCTG-----CCTTCTCCCTTC 1754
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 916 TTGTTGAAGATATTTGCTGTCTCCCTTAAAGTTGAAAAATTTTCAATCTCATTTTCTCCTATT 857
QY 1755 CTCTATCTCTTTGTTTGGCATAGTCTCTGCTTCCCTGATGTTTATGCTGCGATT 1814
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 856 GTCTGTAGTTTGGTCTCTCTATCTGCTGCTGATTTCTGATGTTTGGAGTTAGATC 797
QY 1815 ATTTPAGCTTAACATTTCTTTTGACCAAGGATTCATTTCTTCAATCTCTCTCACTG 1874
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 796 TTTTGGCATTTTGCATTTCTTTGATGTTGTCGCCATGTTCTCTATGAGTTTCTGCA 737
QY 1875 CCTGAGATCTCTCTCTATCTCTGTTATCTGTCAGTGAGGCTGTCTCTGAGTTCCT 1934
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 736 CCTGAGATCTCTCTCTATCTCTGTTATCTGTTGTCGTCGTCACATCTATGTTGCGA 677
QY 1935 G-TTGGGTTCTTAAATTTTTCATTTCCAGATTTCTTCCAGTTTGGGTTTGTATTAAT 1993
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 676 GATTTCTTCTAGGTTTATATCTCCAGCATTTGCTCACITTTGGGT---TTTATTGTT 621
QY 1994 TCTATTTCCACTTTCAGTCTCGGAATGTTTATCTATCTTCTCTCCCTCC---AGTATTACA 2050
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 620 TCTACTTCCCTTTTATGTTCTTTGATGTTTGTTCAGTCCATCCTGTTGATTTG 561
QY 2051 TTTTTCATAGTTTCTTAAATGATTTATTTCTTCTTCTTCAAGGACCTTTTATGAATC 2110
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 560 TTTTCTGTTATTTCTTAAAGATCTTCACTTCTTATCAGTAGTGTCTGTATTTCTTTA 501
QY 2111 ATAAATGTATGTTAAGTCTGCTTGTGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCT 2163
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 AGTGAGTTATTAAAGCCCTTCTTGTATGTCCTTCTTCCAGCATCATGAGATATGA 448

```

RESULT 4

AK033883

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 3235)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inctani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saich, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1. .3235

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:9330109E13"

/db_xref="MGI:2398061"

/db_xref="taxon:10090"

/clone="9330109E13"

/sex="male"

/tissue_type="diencephalon"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

Eukaryota; Metazoa; Chordata; Craniata, Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
9279253
10349636

2
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20493374
11042159

3
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Taghiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawaj, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
10748366

6
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

(bases 1 to 4895)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kotani, K., Kohjima, Y., Kondou, S., Konno, H., Koumoto, M.,
Koyas, S., Kuritara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakatsuma, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saio, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawara, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Takigawa, A., Takahashi, F., Takaku-Akaira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome.res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/ Tel:81-45-503-9222

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.
URL:<http://genome.gsc.riken.go.jp/>
URL:<http://phantom.gsc.riken.go.jp/>

Qy 887 CTTCTTAGGGCAATTTGCTTAGAATA-TCTTTTCCATCCTTTTACTCTAAGGTGATGCT 945
Db 676 TTTCTTCAGACCAATTTGCTTGGAAACCTGTTTCCAGCCCTTCTTCTGAGGTAGTGCT 735
Qy 946 ATCAAT---GGTAGTGTCTTTTGGATGCACAGTAGAGATGATCTGTTTTCATAT 1002
Db 736 CTCGTGTTCCCTGAGATGAGTTTCTGTAAGACGAAAGTTGGGCTGTTTGTGTAG 795
Qy 1003 CCATTCCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCAATGATGATGTTG 1062
Db 796 CCAGTCTGTAGTCTAATGCTTTTATTTGGGAGTTTAGTCCATGA-----TA 844
Qy 1063 AGAATTAATCAATGACAGTGTGTTGGGATCTTGTATCTTGACCTGTGAGGTGTGT 1122
Db 845 AGAGATATTAAGGAAACATAATTTTGGTCTTATTTATTTT-----TGT 888
Qy 1123 GT 1182
Db 889 TGTGAAAGTGGCAATCTGTTCTTGTGCTGTCTTCTTTAGGTGTTGTAAGTTTACT 948
Qy 1183 GT 1242
Db 949 TCTGTGCTTATCTAGGATGTTTCTATCTGTATTTTCTTTTCTGTTATTTAT-- 1006
Qy 1243 ATTATTCATATTTCTTGAATGTGGTAACATCTTAGATTGAAGTTTCTCTAGCCT 1302
Db 1007 -----CC 1008
Qy 1303 TCTTTAGTGTGCAATTTGAAGATAGATATCTTTACATCTGATTTATCTTAGAATGCT 1362
Db 1009 TTTGAAGGACTGGATTCGTGGAAGATAATCTTTGAATTTGTTTGTGTGTAATCTT 1068
Qy 1363 TCTTTTCCAACTATTTGTGACAGAAAGTTTCTTAAGTGAGTAGTGTGCGCTGACATC 1422
Db 1069 TAGATTTCCATCTATTTGTAATGAAGTTTGGCTGGGTATAGTAGGCTGGGCTGCA 1128
Qy 1423 TGTAGTCTCTGGAGTGTGACATCTGTGACGGGCTTCTTACATTTTGTGCTTCTA 1482
Db 1129 TGTGTCTCTTAGTGTGTATACATCTGTCCAGTATCTTCTGCTTCTCATGTTCTG 1188
Qy 1483 TTGGAAGAGTCAGGTGTAATTTCTAATACATCTGCTTTTATATGTTAATGCTTTTTC 1542
Db 1189 GT-GAAAGATCTGGTGAATTTCTGATAGCGCTGCTTTATATGTTACTT-GACCTTTTA 1246
Qy 1543 CTTTGCATCTTTAATATCTTCTTCTGCTCTATCTTATCTTATCTTATGATGATATGCA 1602
Db 1247 CTTTACTGTTTAAATATCTATCTCTTATTTAGTGTCTATTTCTGTTGATATATGAG 1306
Qy 1603 CTGTGGGAGTTCTTTTCCGCTCCAAATCTATTTGGTGTGTTGTATGCTTCTTACCTT 1662
Db 1307 TGGGAAGATTTTCTCTGCTGCTGCTATTTGGAGTCTGTAGGCTTCTGTATGTT 1366
Qy 1663 GATAGGATCTCTTCTCAAGGTAGGAAATTTTCTTTTGGTGTGTTTCTTGAATAATTT 1722
Db 1367 CATGGCATCTCTTCTTAAAGTTTGAAGTTTCTTCAATAATTTTGTGTAAGATATT 1426
Qy 1723 TTCCCTGTTTGGACCTG-----CCTTCTTCCCTTCTCTATCTTCTTCTTCTTGG 1768
Db 1427 TGCTGGCCTTTAAGTTGAATAATCTTCACTCTCTACTCTATCTATCTATCCATAGTTGG 1486
Qy 1769 TTTTGCATAGTGTCTGGCTCTGGAAGTTTATGCGCTGGAATTTTATGACTTAAC 1828
Db 1487 TCTTCTCATTTGCTGGAATTTCTGGAATTTCCAGTTAGGATCTCTTGTGCAATTTGC 1546
Qy 1829 ATTTTCTTTGACCAAGGTATCCATTTCTCTATCTTCTTCTTCTTCTGCTGAGATCTCTC 1888
Db 1547 ATTTTCTTTCTATGTTGTGCTGATGTTCTCTATGGAATCTTCTGACCTGAGATCTCTC 1606
Qy 1889 TTCTATCTCTTGTATTTCTGAGTGTGAGGCTTGTCTCTGAGGTCTCTG-TTGGGTTCTTAA 1947
Db 1607 TTCCATCTCTGTAATCTGTTGCTGATGCTCTCTCATCTATGTTCCAGATTTCTTCTCTAG 1666

Qy 1948 TTTTTCATTTCCAGATTTCTTCAAGTTTGGTGTGTTTGTATTAATTTCTATTTCCACTTT 2007
Db 1667 GATTTCTATCTACATTTGTTGACTCTTCTGGGTTTCTTTTATTTGTTGTTCTACGCTCTTT 1726
Qy 2008 CAGTCTCTGAAATGTTTACTCATTTTCTCCAGTATTTTACA---TTTTCATAGTTT 2063
Db 1727 TAGTCTAGTATGATTTTGTATTTCATTTCCATCTGTTTGGATGTTTCTGTTTTT 1786
Qy 2064 CTTTAATGATTTTATTTCTTCTTCAAGGACCTTTTATGAATTCATAAATGTA 2120
Db 1787 CTTTAAGGACTTCTACCTCTGTTGTTGTTGTTCTCTCTGTTTCTTTTAAGGACTTGA 1843

RESULT 10
AK046811/c
LOCUS
DEFINITION

AK046811 4832 bp mRNA linear HTC 20-SEP-2003
Mus musculus 10 days neonate medulla oblongata cDNA, RIKEN
full-length enriched library, clone:BB30013D20 product:unknown EST,
full insert sequence.

ACCESSION AK046811 GI:26091732
VERSION HTC; CAP trapper.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20493374
PUBMED 11042159

REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE 4
AUTHORS FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED 11076861

REFERENCE 5
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12042159

REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE 7
AUTHORS FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED 11076861

REFERENCE 8
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 12042159

REFERENCE 9
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sueni-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

1..4832
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:B830013D20"
/db_xref="MGI:2412332"
/db_xref="taxon:10090"
/clone="B830013D20"
/tissue_type="medulla oblongata"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="10 days neonate"
1..4832
misc_feature
/note="unknown EST (GB|BM120546, evidence: BLASTN, 100%, match=704)"

Query Match 22.6%; Score 515.998; DB 1; Length 4832;
Best Local Similarity 61.4%; Pred. NO. 0.83;
Matches 1232; Conservative 0; Mismatches 480; Indels 296; Gaps 28;

QY 292 TCTATTTCTGATCTCTATCTGCTCATTTTAACTCAGTAGTGTGTTGGTTCC 351
DB 2612 TTTCTTTCTTAATTTCTGCTAGACCAATATTCATCA-----GTTCAGTTCC 2563
QY 352 ATAAGTTTGAAGTTTCTGTTTCTGTTGTTGTTGTTGTTATCTAGATTAAAGCTGT 411
DB 2562 ATGAGTTTGAAG-----CTTTCAGATTGTTGGTGTATCTAGTATTAACTGT 2512
QY 412 GGTGTCAGATAGGCATAGATATATTTCAATGCTTTTATCTGTCGAGCTTCCT 471
DB 2511 GGTG--CTGACAGGATGAGAAATGTTATTTCAATTTTCTTATATTTGTTAAACATGCTT 2454
QY 472 TGTTTGAATATGATTCA-ATTTGGAGAGTTTCATAGGTCCTGCACAAAGAGTACA 530
DB 2453 TGTGTCAGTATGTGTCAGCTTTTGAAGATTCATGAGTCCTGCAAAAAGAGCAT 2394
QY 531 GT--CTTTGTTTGGTGAATAGTCGTGAAATATCT-CTAGTCCACTGTTGTTATGA 587
DB 2393 ATTCCTTTGTTGTTAGGTGTAATGTTCTGTGACTATCTACTTGTCTCAATTTAGCA 2334
QY 588 CATCAGTTAGTCCAGCATTTCTCTGTTTCTGTTTGTGAGATGACCTAACTGTTGGA 647
DB 2333 CATCATTTAATCCAGCAATTTGCTATTTAG-TTTTGTCTGGAAGGCTGTCTGTTGGT 2275
QY 648 GAGAAATGGGTTTGAAG--TAGCCCACTATCTGTGTGTA--GGTCAATATGTGATTTAG 705
DB 2274 GACAATGGGTTACTGAAGTTTCCCACTATCATCAGTGTATGAGGACATGCAATTTAG 2215
QY 706 CTGTAGCTGTCTGTTTATGAACTTGGGTCATGTTGTTGGTGCATAGACATTAAAG 765
DB 2214 CTGTAGTAGTGTCTTTTTTGAACCTTGGGGCCCTTGT-----ATAAG 2171
QY 766 AATTGCAATGTCCTCTTGGTGGATTT-TCCTTTGATGCTATGATGATTTCTTCCCAATC 824

DB 2170 AACGGAATGTCCTTAGGTGAATTTCTCTTTTGATGAATATGTAAGTCTCTTCCCAATC 2111
QY 825 TCATCTGCTTAGTTTTGGGTTTAAAGTCTA-TTAGTCAGATATTTAAATGACTGTATCGGC 883
DB 2110 TCCTCTAATCAGTTTTGGTCTGAACCTCTATTTTGTGAGATAAGAAATAGCTACAGCAGT 2051
QY 884 TTGCTTTTAGGCCCATTTGCTTAGAATATCTTTTCCATCTCTTTTACTCTAAGGTGATGT 943
DB 2050 TTGCTTTCTTAGTCCATTTGCTTAGAATATCTTTTCAACCTTTTCTGGGGTGATACCTA 1991
QY 944 CTATCCATGTTAGTGTCTTTTTTGGATGCAGCAGTAGGATGATCTTTT-----995
DB 1990 CCCTTGATTTAAGGTATGCTCTTAGATGAAGCCACAGATAGATTTCTTTGTTTCAA 1931
QY 996 -TTCATATCCATTTCTGTTACCCAGTATCTTTTCTAGAGAAATTAAGATCATGTAGTCAT 1054
DB 1930 ATGGATTTTCTGATTTCAAATCCATCTCTGTAGTCTCTCTCTTTTATTTGGGAAC 1871
QY 1055 TGATGTTGAGAAATTAATCAATGAGCAGTGTGTTGTTGATCTCTGTTATCTTGCACTTGTGAA 1114
DB 1870 TGA-GACAAGAGGTACCAACAACAATGTTGTTGACTTCCTGTTATTTGCTACTATAG 1812
QY 1115 GTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1174
DB 1811 GTATATAAG-----1803
QY 1175 GTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1234
DB 1802 -----TTTCCCTCTTTTGGATTTGCTGATCTG 1776
QY 1235 AATTAATTAATTAATTAATTTCTTGAATGTTGGTGAACATCTTTAGATTGAAGTTTCT 1294
DB 1775 GAATAAGTATT-----1765
QY 1295 CCTAGCCTTTTAGGTCTGCAATTTGAAGATAGATATTCTTTACATCTGATTTTACTTA 1354
DB 1764 -----TGGTTTATCAAG 1752
QY 1355 GAATGCTTTTCTTCTCAAACTANTTGTGACAGAAAAGTTTTCTAAGTCAGTAGTCTGCG 1414
DB 1751 GGTCTATTTATTTCTCTCAACTATGACAAATTTGAAAGTTTTGCTAGTATTAATTTGGAG 1692
QY 1415 CTGACATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGACAGGCTCTTTTACATTTTG 1474
DB 1691 CAGGTATCTGCAGGATCTTAGAGTCTGTAGAACATTTGTTGAGGCTCTAGAGAATCTCTA 1632
QY 1475 AGTTTCTATTTGAAAAGTCAGGTGAATCTAATAACATCTGCTTTATATGTTAATTCGT 1534
DB 1631 CTA-----AGAACTCAGGTATTTACTCCAAATTTGGTGTGCTGTATATGTTACTTGGT 1581
QY 1535 CT-----TTTTTCCCTTGCATCTTTTAAATATTTCTTTCTTTGTTCTATACCTTT 1581
DB 1580 CTTTTTCCCTTGGTCTTTTCCCTTGCAGCTTTTAAATTTCTTTCTCTCTGTAGGTT 1521
QY 1582 TAGTGNATTGATATATGCACTGTGGGAGTTTCTTTTCCGGTCCCATCTATTTGGTGT 1641
DB 1520 TAGTGTCTCAATTTATATGTGCCAATAGAGTGTCTTTCTGGTCCAGTCTACTTTGGTG- 1462
QY 1642 TTTGTATGTTCTTCTTACCTTGATAGGCATCTCTTTTCTCAAGGTTTAGGAAATTTTCTTT 1701
DB 1461 -----CTCTATATCTTCCATAGATCTCCATCTTTATGTTAGGAAATTTCTTTC 1411
QY 1702 TTTGGTTTCTT-GAAAATATTTTCCCTGCTTTTGAACC-TGCTTCTTCCCTCCCTCTA 1759
DB 1410 TGTGATTTTTTTGGAAAAATATTTTCTGTGCTTGTCTCTGATTTCTTCTCTTCTCTA 1351
QY 1760 TTTCTTTGGTTTTTGCAT--AGTGTCTCTGGCTTCTCTGGATGTTTATGCTGGAATATT 1817
DB 1350 TTCTCTATTTCTTTAGATCAAGTGTTCAGATTTCTCGAATGTTTATAGAGGATTTT 1291
QY 1818 TTAGACTTAAACATTTTCTTTTGAACAGGTATCCATTTCTTCTATCTTGTCTTCTGCT 1877


```

Db      1290 TTT-----CCTGTTTAAACATCTGAGGTATCCATTTCTTCCATCATGTCTTTAATGCCT 1238
Qy      1878 GAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1936
Db      1237 GAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1178
Qy      1937 TGGGTCTCTAATTTTTCATTTCCAGATTTCCCTTCAGTTTGGGTTTGGTTTATTAATCT 1996
Db      1177 TCCAATCCCTAAATTTTCATTTCCAGATTTCCCTTCAGTTTGGGTTTGGTTTATTAATCT 1118
Qy      1997 ATTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2055
Db      1117 ATTACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1058
Qy      2056 ATAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2115
Db      1057 GATG-----CAAGGATTTTATTCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1004
Qy      2116 ATGATGTTAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2174
Db      1003 GGTATTTTAAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 944
Qy      2175 AATAGGCT-----TTTACAGGACATATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2226
Db      943 TGTAGGTTACTAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 884
Qy      2227 GCTTGGCATATAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2254
Db      883 ATGGCAGCATCTAGGCATCTGGGACTGG 856

RESULT 11
AK045866
LOCUS      AK045866 3375 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus adult male corpora quadrigenima cDNA, RIKEN
full-length enriched library, clone:B230315G06
product:unclassified, full insert sequence.
ACCESSION AK045866
VERSION AK045866.1 GI:26091171
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED 10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
PUBMED 11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Okawa, K., Tanaka, T., Matsubara, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED 11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the

```

```

PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3375)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, K., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fastom.gsc.riken.go.jp/.
FEATURES
Source
Location/Qualifiers
1..3375
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:B230315G06"
/db_xref="MGI:2411246"
/db_xref="taxon:10090"
/clone="B230315G06"
/sex="male"
/tissue_type="corpora quadrigenima"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1..3375
/notes="unclassified"
Query Match 22.1%; Score 503.498; DB 1; Length 3375;
Best Local Similarity 62.4%; Pred. No. 1.3;
Matches 1091; Conservative 0; Mismatches 490; Indels 167; Gaps 21;
Qy 384 TTGTTGTTGTTATCTAGATTAAAGTCTGCTGTCAGATGAGCATAGATGATTTATTTCA 443
Db 1 TAGTTGTTGAAGCCAGCTTTAATCCATGCTGATCTGATAAAACACACGGGGTCATTCA 60
Qy 444 ATTGTTCTTTTATCTGTCGAGACTGCTGTTTGTGTTTGAATATGTTCAATTTTGGAGAGT 503
Db 61 ATCTTTCTGTAUCTTTTGGAGCTTACTCTGCTGATCATATGTTCAATTTTAGAGAAG 120
Qy 504 TTCATAGG--GTGCTGACAGAGGTCAGTC--TTTGTTGTTTGGTGAATAGTCGTGTA 560
Db 121 ATTCTGGGATGTAAGGAGGAGGATATCTTTTGTGTTGTTGAGTGAATGTTCTGCA 180
Qy 561 ATATCTC-TAGGTCCACTTGGTTTATGACATGAGTTAGCTCCAGCATTTCTCTGTTTCT 619
Db 181 ATATCTCTAAGTCCACTTGGTTTATGACATGAGTTAGCTCCAGCATTTCTCTGTTTCT 237

```

QY 620 TTTTGTGATGACCTAACTGTT--GGAGAGAAATGGGGTAATGAAGTAGCCCACTATC 677
 Db 238 CTTCTGTTTATGATGACCTGTCATTTGAGGTGAGAGTGGGGTGTTGAAGTCTCCCAATATT 297
 QY 678 TGTGTGTGAGG--TCAATATGTGATTTTAGCTGTAGCTGTGCTGTGTTTATGAACCTGGGT 736
 Db 298 ATTGTGTGAGGTCAATGTGTGATTTGTGCTTTAGTAAAGTTTCTTTTACAAGTGTGAGT 357
 QY 737 GACATGTGTTTGTGTCATAGACATTAAGAAATGAATGTCCTCTGTGTGATTTTCCCTT 796
 Db 358 GGCCTTGCATTTGAGTACAGATGTTTCAAGATGAGATATCATCTTGTGTAGATTTTCCCTT 417
 QY 797 TGATGCCCTATGTATTTTCCCAATCTCATCTGTCTAGTTTGGGTGTTAAGTCTA-TT 855
 Db 418 TGATGCATATGAAGTGTCTTCTCACTCTTTTGAATCTTTTGTGTGTTGAAATCCATTT 477
 QY 856 AGTCAGATATTAATAATGACTGTATGGCTTCTTTAGGGCCATTTTGTCTAGAAATATC- 914
 Db 478 TATTAAATATTAGATTCACTACTCCAGCTTGTCTTCTGGGTGTTGGCTGGAAACCT 537
 QY 915 TTTTCCATCCTTTTACTCTAAGGTGATGCTA---TCCATGTAGGTGTCTTTTGGGA 971
 Db 538 TTTTCCAGCCCTCTATGCTGAGGTAGTGTCAACTTTGTGTGCTGAGATGTTTCTTTGTA 597
 QY 972 TGCAGCAGTAGGATGGATCTTGTGTTTTCATATCCATCTCTGTACCAGATATCTTTTCTAG 1031
 Db 598 TGCAGCAGATGATAGATTCTGTTTGTGTATCCAGTCTTTTAGCTGTGCTTTTATCA 657
 QY 1032 AGAAATTAAGATCAATGAGTCATGATCTTGAGAAATTAATCAATGAGCAGGTGTTGTGGAT 1091
 Db 658 GAGAACTGAGTCCCACTGA---TGTTGAGATGAGAGATATTAATGATCAGTGTGTTGTATT 714
 QY 1092 TCTGTGATCTTGCACTTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1151
 Db 715 TCTGTGATTTTGTGTTG----- 733
 QY 1152 TGT 1211
 Db 734 -----GGTATGT 755
 QY 1212 CCTCTTTTGTGTTTGGCTGGATATTTTATTTATTCATATTTCTTGAATGTGGTAA 1271
 Db 756 -----TCTTTTCTGTGAATGAGTAATTTCTTGTGTCTTCTTGGGGGTAGTTAC 804
 QY 1272 CATCTTTAGATTGAAGTTTTTCTCTAGCCCTTTTAGGTCTGCATTTGAAGATAGATAT 1331
 Db 805 CCTCCTTGTGTGGAT-----GTAGATAT 828
 QY 1332 TCTTTACATCTGATTTTATCTAGAAATGCTTCTTCTTCTCAACTAT-TGTGACAGAAAG 1390
 Db 829 TGTTTAACTTTGTTTATCATGAGGATATTTTGGTTTCTCCATCTATCAGTGAATTAAGAG 888
 QY 1391 TTTTCTAAAGTCAGTACTGTGGCTGCATCTGTAGTC--TCTTGAGTCTGTAGACACA 1448
 Db 889 TCTTGTGTTGTTATAGTACTGTGGCTGCCATCTGTCTTTTTAGGCTGTAGACACA 948
 QY 1449 TCTGTGAGGCGCTTCTTACATTTTGTAGTTTCTATTTGAAAGTCAAGGTGAATCTTAAT 1508
 Db 949 TCTGTTACAGATCTTCTGGCTTTTATATAGTCTCTGTGAGAAATCAG---GTAGTCTGTAT 1005
 QY 1509 ACATCTGCTTTATATGTTAATGCTTCTTTTCTCCCTTGCATCTTTTAAATATCTTTCTT 1568
 Db 1006 AAGTCTGCTTTTATAGTGTCTTGTGGCTTTTCTCTGTGAGCTTTTAAATATCTTTCTT 1065
 QY 1569 TGTCTTATPACTTTTAGTGATTTGATTAATGATGCTGTGGGAGTTTCTTTTCGGTCCA 1628
 Db 1066 TGTCTGTACATTTAGTGTGTTTGTATTAATGTCGAGGAAGATTTTCTTTTGTGCTA 1125
 QY 1629 ATCTATTGTTGTTTGTATGCTTCTGTACCTTGTAGCAGTCTCTTCTCAGGTGTAG 1688
 Db 1126 ATCTATTGTTGTTCTGTATGCTTCTCATGAGTGTGTTTGTAGTATCTCTTT-TTTAGGTGTG 1194

QY 1689 GAAATTTTCTTTTGGTTTCTTGAAATATTTTCCCTGCTTTTGACCTG----- 1740
 Db 1185 GGAAGTGTCTTCTATGATTTTGTGAAGATATGTTCTGTCCCTTTGAACCTGGGATCTT 1244
 QY 1741 -----CCTTCTTCCCTTCTCTATTTCCCTTGGTTTGTTCATAGTGTCCCAAATTTCCCTG 1795
 Db 1245 CACCTTCTTCTATCTTATTTCTTAGATTTGGTCTTTTCTATAGTGTCCCAAATTTCCCTG 1304
 QY 1796 GATGTTTATGCTGAGTATTTTATGACATTTCTTTGACCAGGATTCATTTTCC 1855
 Db 1305 GATGTTTCTGAGTTAGAACACTTTTACATTTGGCATTTTCTTTGACATATCTATCAATTTA 1364
 QY 1856 TTTATCTTGTCTTCACTGCCCTGAGATCTCTCTTCTATCTTCTTGTATTTCTGTCTAGTCAG 1915
 Db 1365 TACTATGGGATCTTCTACACCTGAGATCTCTCTTC-ATCTCTTCTATTTGTGTGTAAT 1423
 QY 1916 GCTGTCTCTGAGGTCTCTGTGGGT--CTTAATTTTCTATTTCCAGATTTCTCTCAG 1973
 Db 1424 GATTCATTTACTAGTCTCTTCTTCTTCTAGATTTTCCAACTCCAGGTTGCTCCA 1483
 QY 1974 TTTGGTGTGTTTATTAATTTCTATTTCCACCTTTCAGTCTGAAATGTTTACTCATTT 2033
 Db 1484 TTCAATGTTTCTCATGCTGCTTCTATTTCTGTTTCTGAGTCTTGAACAATTTTATCATTT 1543
 QY 2034 TCTCCAGTATTTACATTTTCTATAGTCTTCTTAAATGATTTATTCATTTCTCTCAA 2093
 Db 1544 CTTTACCTGTTGATTTGATCTTTTCTATTTCTTTAAGGATTTCTTCTCTCTTAA 1603
 QY 2094 GGACCTTT 2101
 Db 1604 GGACTTCT 1611

RESULT 12
 AK034168
 LOCUS
 DEFINITION
 Mus musculus adult male diencephalon cDNA, RIKEN full-length
 enriched library, clone:9330161A16 product:unclassified, full
 insert sequence.
 ACCESSION
 AK034168
 VERSION
 AK034168.1 GI:26083804
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 PUBMED
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 PUBMED
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 PUBMED
 11076861

Query Match		20.1%; Score 457.898; DB 1; Length 2108;
Best Local Similarity		62.0%; Pred. No. 3;
Matches 1034; Conservative		0; Mismatches 496; Indels 137; Gaps 21;
QY	482	TAATGATTCATTTTGGAGA-GTTTCATAGGGTGTGACAGAGGTACAGTCT-TTGT 538
DB	1973	TATATGTCATTTTGAAGAAGGTTCCATGAGGTGTTGAGAAGAAGATATATTTCTATGT 1914
QY	539	GTTTTGTGAAATAGTCTGTAAATATCT-CTAGTCCACTTGGTTTATAGCATCAGTTAG 597
DB	1913	TTTAGTATAAAGGTTCTAGAAATATCTGTGTAAATCCATTTGGTTTCAACTTCTATTAC 1854
QY	598	CTCAGCATTTCTCTGTTTCTGTTTTTGTGAGANGACCTAACCTGTTGGAGAGAAATGGGG 657
DB	1853	TTTCACTGTGCTCTGTATAGTTTC-TGTTTCCAGGATCAGTCTCTTGTGACAGAGTGGGA 1795
QY	658	TATTTGAAGTACCCACATCTGTGTGTG-TAGGTCATATGTATTTTAGCTGTAGCTGTG 716
DB	1794	TGTTGAAGTCTCCACATATATTTGTGTATAGTCAATGTGTGCTTTTGGAGCTTTACTAAAA 1735
QY	717	CTTGTGTTATGAATCTGGGTGACATTTGTG-TTGGTGATAGACATTAAGA 766
DB	1734	TTTCTTTAAGAAATGTGATGCCCTTGCATTAATAAATATCTTGGAGCATAGATATTCAGA 1675
QY	767	ATTGCAATGCTCTGTGTGATTTTCTTTGTAGTCCCTATGTAGTATTTCTTCCAAATCTC 826
DB	1674	ATTGAGAGTTACCTTAGTAGATTTTACCTTTGATGAGTATGAAGTATCTTCTCTTATCTT 1615
QY	827	ATCTGCTTAGTTTGGGTTTAAAGTCTA-TTAGTCCAGATATTAATAATGACTGTATCGGCTT 895
DB	1614	TTTTGATAACTTTAGTTGAAGTCGATTTTATTCATATTAGATGAGTACTCCAGCTT 1555
QY	896	GCCTTTAGGCGCAATTTGCTTAGAATATC-TTTTCCATCCCTTTTCTCTAAGTGTATGTC 944
DB	1554	GTTTCTTGGGACCAATTTGCTTGGAAATGTTTTCCAGCCCTTTTACTCTGAGGTAGTGTG 1495
QY	945	TATCCATGGTA-GGTTGCTCTTTTGGATCCAGCAGTAGAGTATCTGTTTTCATAT 1002
DB	1494	TGCTTTGCTACTGAGTGGGTTTCTGTATGACAGAAATGCTGGGCTGTTTATGTATC 1435
QY	1003	CCATTTCTTACCAGTATCTTTTCTAGAGAAATTAAGATCAATGATGATGATGTTG 1062
DB	1434	CCAGTCTTTTAAATATATCTTTT-----TATTAGGAAATGAGTCCATGATATTA 1382
QY	1063	AGAAATATCAAGACGAGTGTGTGGAAATCTTGTATCTTGCACTTGTGAAGTGTGTG 1122
DB	1381	AGAGATATAAGTAAAGTAATATTTGTTCTCTGTATTTTGTGTAGAGTTGGAAT 1322
QY	1123	GT 1182
DB	1321	CTGTTCAATGGCTATCTCTTTTAGTTTG----- 1291
QY	1183	GTCTGTGTCTGTGTGTGTGTGTCTCTCCCTCTTTTGATTTTGGCCTGGAATATTT 1242
DB	1290	-----TTGAAGATTA 1280
QY	1243	ATTATTCATATTTCTTGAATGGGTAAACATCTTTAGATTTAGATTTTCTCTAGCCT 1302
DB	1279	TTTCTGTCTTTTCTAGGCAATAGTTTCCCTCTCTTGTGTGATTTTCCATTTATGT 1220
QY	1303	TCTTT-----AGTCTGCAATTTCAAGATAGATATCTTTTACATCTCAATTTATCTAGAAT 1358
DB	1219	CTTTTGAAGGCTGACCTATGAAAGATATATTTTAAATTTGTTTGTCTATGGAAT 1160
QY	1359	GTCTTTCTTCTCCAACTATTTGTGTGACAGAAAGTTTTTCTTAAGTGCAGTGTCTGGCTGA 1418
DB	1159	ACCTTAGTTTCTCCATCTATGTATTTGAGAGTTTTTGTGGGTATAGTAGCTAGGCTGG 1100
QY	1419	CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGACGGGCTTCTTACATTTTGTAGTT 1478
DB	1099	CATTTGTCTCTTCTTAGGTCTGTATGACATCTGCCAGGATCTTCT--GGTTTCTATAA 1042
QY	1479	TCTATTGGAAGAGTCAGGTGTAATTTCTAATAACATCTGCCTTTATATGTTAATGTGCTTT 1538
DB	1041	TCCTCTGGCAAGAAGTTTGGTGTAAATTTGATAGTCTGCTCTTATATGTACTTCACCTTT 982
QY	1539	TTTCCCTTCATCTTTTAAATATCTTCTTCTCTTCTATACCTTTTAGTATGATTATTA 1598
DB	981	TTT-CCTTACTCTCTTTAAATCTATCTTGTGTTGCAATGCTGTTTGTATTA 923
QY	1599	TGCACCTGTGGGAGATTCTTTTCCGGTCCAATCTAATTTGGTGTGTTGTGTGTGTTTGTGTA 1658
DB	922	TTGACAG-GAGGAATTTCTTTTCTCTCTCAATCTATTTGGAGTCTATAGGTTTCTTGTGA 864
QY	1659	CTTTGATAGGATCTTTTCTCAAGTTAGGAATTTTCTTTTGGTGTGTTTCTTGTGAAA 1718
DB	863	TGTTCAATGGCAATCTTTTTCAGGTAGGAAAGTTTCTTCTCTAAATTTTGTGTGAAGA 804
QY	1719	TATTTTCCCTGCTTTTGACCT-GCCTTCTTCCCTCTCTCTATTC-----T 1764
DB	803	TATTTTACTGGCCCTTTAAGTTGGAATCTTCACTGTCTTCTATACCTACTATACAGT 744
QY	1765	TTGGTTTTTGGCATAGTCTCTGGCTCTCTGAGATTTTATGCCCTGGATTTTATAG-AC 1823
DB	743	TTGCTCTTCTCATTTGTCTCGAGTCTCTGATGCTTTGAGTTAGGATCATTTTGTGAT 684
QY	1824	TTAAACATTTTCTTGACCAAGGTATCCATTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1883
DB	683	TTTGCATTTTCTTGATGCTGTGCAAGTTTCTTGGTATCTTCTGATCTGAGAT 624
QY	1884	CT 1942
DB	623	CT 564
QY	1943	CTTAATTTTCT 1998
DB	563	CTTAGATTTTCT 504
QY	1999	TTCCACTTTTCT 2054
DB	503	TTCCATTTTCT 444
QY	2055	CATAGTTTCT 2101
DB	443	CTGTAATTTCT 397
RESULT 14		
AK048889		3179 bp mRNA linear HTC 20-SEP-2003
LOCUS		Mus musculus 0 day neonate cerebellum cDNA, RIKEN full-length
DEFINITION		enriched library, clone:C230078P05 product:carbonic anhydrase 5a,
		mitochondrial, full insert sequence.
ACCESSION		AK048889
VERSION		AK048889.1 GI:26093182
KEYWORDS		HTC; CAP trapper.
SOURCE		Mus musculus (house mouse)
ORGANISM		Mus musculus
REFERENCE		1 Carninci, P. and Hayashizaki, Y.
AUTHORS		High-efficiency full-length cDNA cloning
TITLE		Meth. Enzymol. 303, 19-44 (1999)
JOURNAL		99279253
MEDLINE		10349636
PUBMED		
REFERENCE		2
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE		Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL		Normalization and subtraction of cap-trapper-selected cDNAs to
MEDLINE		prepare full-length cDNA libraries for rapid discovery of new genes
PUBMED		Genome Res. 10 (10), 1617-1630 (2000)
REFERENCE		3
AUTHORS		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

match=663) "

Query Match 18.9%; Score 430.098; DB 1; Length 4502;
 Best Local Similarity 58.4%; Pred. No. 1.9;
 Matches 1038; Conservative 0; Mismatches 579; Indels 161; Gaps 18;

QY 532 TCTTTGCTTTTGGTGAATAGCTCTGTAATAATCT-CTAGGTCACTGCTGTTTATGACAT 590
 DB 5 TTTTITTTTTTAAAGATAAATGTTCTATAGATATCTGTTAAATCCATTTTGTCCATAACTT 64

QY 591 CAGTTAGCTCCAGCATTTCTCTGTTTCGTTTGTGTTGTTGAGATGACCTAACTGTTGGAGAG 650
 DB 65 ATGTTAGTTTCACTGTCCTTTGTTTGTAG-TTTCTGTTTCCATGATCTGTTTCATGTTGCTGAG 123

QY 651 AATGGGTATTGAAGTAGCCCA---CTATCTGTGTGTGAGGT-CAATPATGTGATTTTATGAC 706
 DB 124 GGTGGG-TGTTGAAGTCTCCAAATGTTTATTAATCTGTGTGAGTCAATGTTGTGCTTTGAGC 182

QY 707 TGTAGCTGCTGCTTTTATGAATCTGGGTGACATGTTGTTGGTGCATAGACATTAAGA 766
 DB 183 TTCAGTAAAGTTTCTTTTATGAATGTTGTGCTCCCTGCAATTTGGAGCACAGATGTTTCA 242

QY 767 ATTGCAATGCTCTCTGGTGATTT-TCCTTTGATGCTATGATGATTTCTTCCCAATCT 825
 DB 243 ATTCAGAGTTCTCCTTGGTAGATTCTCCATGATGATGATGAAGTGTCTTCTTACCT 302

QY 826 CATCTGCTTAGTTTGGTTTAACTCTATTAGTCAGATATTAATGACATGATCTGGCTT 885
 DB 303 TTTTGATAACTTTTGGTTGAAGTTGCTTTTATTCGATATTAAGATAGCTGCTCCAGCTT 362

QY 886 GTTCTTTAGGCCCAATTTGCTTTAGAATA-TCTTTTCCATCTCTTACTCTAAGGTGATGTC 944
 DB 363 GTTCTTCCAGACCGTTTGTCTGGAATTTGTTGCCAATGTTTATTTCTGAGTAAATGCC 422

QY 945 TATCCATGGTAG---GTTGCTTTTTTGGAGCAGCAGTAGGATGGAATCTGTTTTCATA 1001
 DB 423 TGTTTTTGTCACTGAGGTGATTTCTCTGTATGAGCAAAATGTTGGTCTCTTTATGTA 482

QY 1002 TCCATTTCTGTACCCAGTATCTTTTCTAGAGAAATTAAGATCATTGAGTCATTGATGTT 1061
 DB 483 TCCAGTCTGTTATCTAATTTCTTTTATTTGGAATTA---ATCCACTGATATT 534

QY 1062 GAGAAATTAACAAGCAGCTGTTTGTGGATCTTGTATTCTGACATCTGGAAGTGTG 1121
 DB 535 AAGAAATTAATAAGAAATATTGTTGCTTCTGTTATTGTTGATGTTATTTTATATT 594

QY 1122 TG 1181
 DB 595 TCTGTGTGTATCTTTTGGGTTGTGTAAGAAAG----- 630

QY 1182 TGTCTGT 1241
 DB 631 -----ATT 633

QY 1242 TATTATTCAATATTTCTTGAATGGGTGAACATCTTTAGATTGAAGTTTTCCTTAGCC 1301
 DB 634 ACTTCTAAGTTTCTTGAAGGTGAGTTTCCCTCTGTGTGTGTGTGTGTGTGTGTGTGT 693

QY 1302 TTCTT--TAGTCTGCAATTTGAAGATAGATATTTCTTTTACATCTGATTTTATCTTGAAGT 1359
 DB 694 TCCTTTGTAGGTTGGAATTTCTGGAAGATATTGTGTGCAATTTGTTTGTGATGAATA 753

QY 1360 TCTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAACTGAGTCTGAGTCTGGCCTGAC 1419
 DB 754 TCTTAGTTTCTCCATCTATGTAATTTGAGAGTTTGTGCTGGTATAGTAGCTGGGCTGAC 813

QY 1420 ATCTGTAGTCTCTGGAGTCTGTAGACATCTGTGAGGSCCTTCTTACATTTTGTAGTTT 1479
 DB 814 ATTTGTGTTCTTGAAGGTTTGTATGACATCTTCCAGGATCTTCTAGCTTTTCATAGTCT 873

QY 1480 CTAATGGAAAGTCAAGTGTAAATTTCTAATACATCTGCTTTTATATGTTAATGGTCTTTT 1539
 DB 874 CTGTGTG-----AGAAGTCTGATAGGACTGCTTTTATATGTTACTT-GACCTTT 920

QY 1540 TTCCCTTGCATCTTTTAAATATCTTTCTTTGTTCTATATCTTTTAGTGATTTGATTTAT 1599
 DB 921 TTCCCTTACTGCTTTTAAATATCTTTCTTTGTTCTATATCTTTTAGTGATTTGATTTAT 980

QY 1600 GCATGTGTGGGAGTTCTTTTCCGGTCCAACTATTGTTGGTGTGTTGATGCTTCTGTAC 1659
 DB 981 GTGTGAGGAGGAATTTCTTTCTGTCTCCACTATTTGGAATTTCTGTAGGCTCTTATAT 1040

QY 1660 CTTGATAGGCATCTCTTTCTCAAGGTTAGGAAATTTTCTTTTGGTGTGTTTCTTGAATA 1719
 DB 1041 GTTTATGGGTATCTTTCTTTAGGTAGAGAAGTTTCTTCTATATAATTTTGTGTAAGAT 1100

QY 1720 ATTTTCCCTGCTTTTGACCT-GCCTTCTTCCCTTCTCTATTC-----TT 1765
 DB 1101 ATTTACTGCTCTTTAAGTTAGGAATCTTCACTATTTCTATATACCTATTTATCTTAGGTT 1160

QY 1766 TGGTTTTTGCATAGTCTCTGCTGCTCTGATGTTTATGCTGGATTTATGCTGGATTTTAGACTT 1825
 DB 1161 TAGTCTTCTCAITTTGCTGATTTTCTGAGTTTGGGTAGGAGCTTTTGCATTT 1220

QY 1826 AACATTTTCTTTGACCAAGGTATCAATTTCTTATCTTGTCTTCACTGCTGAGATCT 1885
 DB 1221 TGCATTTTCTTTGACTGTTGTCTCAATGTTTCTTATGATATCTTCTGGACCTGAGATCT 1280

QY 1886 CTCTTCTATCTTGTATTCTGCTGAGGCTTGTCTCTGAGGTTCTCTGAGTTCTCTGTTGGG-TTCT 1944
 DB 1281 CTCTTCTATCTTATATCTGTTGGATGTTTGTGCAATGACTCTGATCTCTTCTTCC 1340

QY 1945 TAAATTTTCTATTTCCAGATTTCTTCAAGTTTGGGTTTGTTTTATTAATCTTATTTCCAC 2004
 DB 1341 TAGGTTTCTTAACTCCAGGTTGCTCTCTTGTGATAGTTTGTCTCAATCTCTTCACT 1400

QY 2005 TTTCAGTCTCTGAATGT-----TTTACTCATTTTC 2035
 DB 1401 GTTGTGTTGTTTCTCTGCAATTTTAAAGATATTTTGTGTCTCTCTTTAAGGCTTCA 1460

QY 2036 CTCCAGCATTTTACATTTTCTAGGTTTCTTAAATGATTTATTTCAATTTCTCTTCAAGG 2095
 DB 1461 AGCTTTTATCTGATTTCTCTATATCTTTTAAAGGATTTATTTATGCTCTTTGTTAAAG 1520

QY 2096 ACCTTTTATGAATTCATAAAATGATTTAAGTCTCTTGCCTTGTGCTTCACTGATGTTG 2155
 DB 1521 TCCCTCTCATCATCATGAGATATGATTTTAAATCAGAGTCTCTTCTTCTGAGGATCC 1580

QY 2156 CATTTCTCA-GGGCTATTGTAATAGGGTTTGTAGCAGGACATATTGCTGCTGTTGTTATT 2214
 DB 1581 AGGGCTCACTGTTGGGAGAACTGGGTTCTGATATGCCAAGTAGGCTTGGTTCTGTT 1640

QY 2215 GTCTGTGTTTGTCTTTGGCATATAGACGGCTGAGTTT 2252
 DB 1641 GCTTATGTTATTGCTCTTGCCTTTAGACATCTGTTAT 1678

RESULT 16
 BC021486/c
 LOCUS BC021486 2278 bp mRNA linear HTC 04-MAR-2003
 DEFINITION Mus musculus, clone IMAGE:534624, mRNA.
 ACCESSION BC021486
 VERSION BC021486.1 GI:20987654
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2278)
 AUTHORS Strausberg, R.
 DIRECT SUBMISSION
 TITLE Submitted (14-JAN-2002) National Institutes of Health, Mammalian
 JOURNAL Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk
 Email: ccapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NIHSC),
 Gaithersburg, Maryland
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hghri.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 55 Row: c Column: 19
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis

This clone has the following problem: retained intron.

FEATURES

source

Location/Qualifiers
 1. .2278
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGES:5344624"
 /tissue_type="Mammary tumor, C3(4)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI CGAP_Mam6"
 /lab_host="DH10B"
 /note="vector: pcwv-SPORT6"

Query Match 18.8%; Score 428.398; DB 1; Length 2278;
 Best Local Similarity 59.0%; Pred. No. 3.7;
 Matches 954; Conservative 0; Mismatches 546; Indels 116; Gaps 14;

Qy	285	AAACACTCTATTTCTTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTGAGTTGTTT	344
Db	1927	AAAGCTTTAAATTTCTTTATTTCTCTCTGACCAAGTATCACTGAGTA-----TTGTTT	1873
Qy	345	GGTTTCCATAAGTTGTAAAGTTTCTGTGTGTTCTGTGTTGTTGTTGTTATCTAGATT	404
Db	1872	AGCTTCCACATGATGCGGCTTGATTTATTTATGTTGTTATGATATC---AGCCTT	1816
Qy	405	AAGCTGTGGTGGTCAGATAGACATAGATATTTCAATTTGCTTTTATCTGTGAGA	464
Db	1815	AGTCCATGGGATCTGATAGATGATGATGGATTTTCAACATTTCTATATCAGTTGAAG	1756
Qy	465	CTTCTTTGTTTTCGAATATGATATCAATTTGGAGAGTTTCATAGGTCCTGACAGAA	524
Db	1755	CCTGCT-----GATTATATGTCATTTTGGAGAAATGTTGAGAAATGTTTCA	1705
Qy	525	GGTACAGCTTTTGTGTTTGGTGAATAGTCTGTGTAATATCTCTAGGTCCACTTGGTTA	584
Db	1704	GGTATATCTTTCTTTTAGGATAGATGTTCTATAGATATGTTATCATCATTTGTTTCA	1645
Qy	585	TGACATCAGTAGTCTCAGCATTTCTCTGTTCTGTTTCTTTTGTGAGATGACCTAATCTGTT	644
Db	1644	TAACTTCTGTAGTGTCAATGTATGTTG-TTGAATTTGTTTTCAGGATCTGTCCACT	1586
Qy	645	GGAGAGAAATGGGTATTTGAAGTAGCCCACTATCTGTGTGTG-AGGTCAATATGTGATTTT	703
Db	1585	GATGAGATGGGAGTTGAAGGCTCCCACTATTTATGTTGAGCAGTGCATGTTGTTTGG	1526
Qy	704	AGCTGTAGTGTGCTTGTGTTTATGAATCTGGGTGACATTTGTTTGGTGCATGACATTA	763

Db	1525	AGCTTTACTAAAGTTTCTTTAATGAATGTGGATGCCCTTGCAATTTGGAGCATATATGTTTC	1466
Qy	764	AGAAATGCAATGTCTCTTGGTGGATTTT-CCTTTGATGCCCTATGTAGTATTTCTTCCCAA	822
Db	1465	AGAACTGAGAGTTCACCTTGGTAGATTTTACCTTTTGATGAGTATGAAATGCCCTCCCTGT	1406
Qy	823	TCTCATCTCTCTAGTTTGGGTTTAAAGTCTATTTAGTCAGATATTAATAATGACTGTATCGG	882
Db	1405	CCTGTTCTTTTGGTAACTTTAGTTTAAAGTCAAATTTATTTCTATATCTATCTACTCCAG	1346
Qy	883	CTTGCTTTCTTAGGGCAATTTGCTTAG-AATATCTTTTTCATCTCTTTTACTCTAAGGTGAT	941
Db	1345	CTTGTTTCTTTGGGACCATGTGCTTGGAAAAATGTTTCCCAACCTTTTACTCTAAGGTAGC	1286
Qy	942	GTCTATCCATG--GTAGTGTCTCTTTTGGATGCAGCAGTAGGATGATCTTGTGTTTC	998
Db	1285	ATTGTCCTTTGACCTGAGTAGGTTTCTGTATGACGCAAAATGTTGAGTCTCTGTTTAC	1226
Qy	999	ATATCAATTCCTTACCAGTATCTTTTCTTAGAGAAATTAAGATCATTTGAGTCAATTGAT	1058
Db	1225	AAAGCCAGTCTGTAGTCTATGCTCTTTTATTGGTGAATTTGA-----ATCCATTGAT	1174
Qy	1059	GTTGAGATTAATCAATGAGCAGTGTGTTGGATTTCTTGTATCTTCTGCTTCTGCACTTGTGAGTGT	1118
Db	1173	ATTGAATGATATTAAGGAAAGTAATTTGTTGCTTCTCTGTTATTTT-----	1129
Qy	1119	GT	1178
Db	1128	-----	1129
Qy	1179	CTGTGTGTGTGTGTGTGTGTGTGTCTCTCCCTCTTTTGATTTTGGCCTGGAATT	1238
Db	1128	-TTGTTGTAGATTTTGAATTTCTGTTTCCATTTGGCTATCTTCTTTTGGTGGTGGTGA	1070
Qy	1239	ATTATATTTATTCATATTTCTTGAATGTGGGTAAACATCTTTAGATTTGAAATTTTCTCTCTA	1298
Db	1069	TTATTTCTTGTCTTTTCTAGTGTGTGTTTCCCTCTCTGTTGCTGTAGTTCCTCTTAT	1010
Qy	1299	GCCTCTTTTGTGTGCAATTTGAAGATAGATATCTTTACATCTCATTTTATCTTAGAAT	1358
Db	1009	ATCCTTTGATGGCTGGATTCATGTAAAGATATCTGTAAATTTGCTTTTGTTCATCGGAAT	950
Qy	1359	GTCTTCTTTTCCAACTATTTGTGACAGAAAGTTTCTTAAGTCAGTAGTCTGCGCTGA	1418
Db	949	ATTTTGTTTCTCCATCTAAGTAAATTTGAGAGTTTGTGAGTATAGTAGCTGGCTGG	890
Qy	1419	CATCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGCGAGGCTTCTTACATTTTGAAGTT	1478
Db	889	CATTTGTAATCTTTAGGCTATATATGACATCTGTCCAGGATCTTCTGGCTTTTATAGTC	830
Qy	1479	TCTATGGAAGATCAGGTGTAATTTCTAATACATCTGCTTTATATGTTAATTTGTCCTTT	1538
Db	829	TCCGGT-GAGAAAGTCTGTTAATTAATAGGCTGCTGCTTCAATCTTACTT-GACCTT	772
Qy	1539	TTTCCCTTGCATCTTTTAAATTTCTTTCTTCTTCTATCTTTTAGTGAATTTGATTATTA	1598
Db	771	TTTCCCTTACTCTTTTAAATTTCTTTCTTGTGTGTCATTTTGTATTTTGAATATTA	712
Qy	1599	TGACCTGTGGGAGTTTCTTTTCCGGTCCAAATCTATTTGGTGTGTTTGTATGTTCTTGTGA	1658
Db	711	TGTGATGGGAAGAAATTTTCTTCCAGTGCAAACTATTTGGAGTCTCTGTAGGCTCTTTATA	652
Qy	1659	CCTTGATAGGCATCTCTTTCTCAAGTTAGGAAATTTTCTTTTGTGTTTCTTCTTGA	1718
Db	651	TGTTCAAGAGAAATCTTTTCTTTAGGTAAGAAATTTTCTTCAATTTTGTGTAAGA	592
Qy	1719	TATTTTCCCTGCTTTTGAACCTG-----CCTTCTTCCCTCTCTATTTCTTTGGT	1772
Db	591	TATTTTACTGGTCTTTTAAAGTGAATCTTCAATTTTCACTTATTTGCTTAGGTTTCTTCT	532
Qy	1773	TGATAGTGTCTCTGGCTTCTGGAATGTTTATGCTGCTGGAATTTTATTTAGCATTAACATTT	1832
Db	531	TTCAATGTGTCTCTGGAATTTCTGGAATGTTTGAAGTAGGATCTTTTGCATTTTGCATTT	472

[illegible]

/db_xref="MG1:2418872"
 /db_xref="taxon:10090"
 /clone="D030041121"
 /tissue_type="whole body"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="9 days embryo"
 1407..1748
 CDS
 /note="unnamed protein product; putative
 similar to ORF 1 [Mus musculus] (SPTR|O61786, evidence:
 FASTV, 71.7%ID, 98.5%length, match=821)"
 /codon_start=1
 /protein_id="BAC34469.1"
 /db_xref="GI:26341614"
 /translation="MEIETIKKSQRETLSEIENLGRSGVIDASITNRIQIEKERIPG
 AEALENIDTAVENAKRCKLLQNIQEIQKWRPNLWIGIEEDSQLKGPVDIFNK
 IIEENPHNLKK"

Query Match 18.1%; Score 412.599; DB 1; Length 2535;
 Best Local Similarity 59.8%; Pred. No. 3.8;
 Matches 851; Conservative 0; Mismatches 424; Indels 148; Gaps 11;
 798 GATCCCTATGATGATATCTCCCAATCTCATCTGCTTAGTCTTTGGGTTTAAAGTCTA-TTA 856
 Db GCTGAGTATGAAGTGTCTCTCCCTATCTATTTTGATAACTTTTGGTTGAAGTCAATTTT 2476
 857 GTCAGATATTAAATGACTATACGGCTGCTCTTAGGGCCATTTCTTAGAATA-TCT 915
 Db ATTTGATATTAGAAATGTTATCTCAACTGTTTCTTGAGACCAATTTGATTGGAATTTGT 2416
 916 TTTCATCTCTTACTCTAAGGTGATGCTATCCATG---GTAGGTGTCTTTTGGAT 972
 Db TTTCAGCCCTTACACTGGGAGTGTCTCTTGTCTGCTAGGTGAGTTCCTGTGAC 2356
 973 GCACAGTAGGATGGATCTTGTGTTTCATATCCATCTGTTACCAAGTATCTTTTCTAGA 1032
 Db ACACGAAATGTTGGGCTTATTAATGATCCAGTCTGTAGTTATTTCTTTTGGAG 2296
 1033 GAAATTAAGATCATTGATGATGATCTTGGATATCAATGAGCATGTTTGTGGATT 1092
 Db GGAATGAG-----TCCATGATATTAAGATATTAAGAAAGTATGATGATTCCT 2244
 1093 CTTGTTATCTTGCACTTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1152
 Db CTTGTTATCTGTGTTAGAGT 2188
 1153 GTCTCT 1212
 Db ----- 2188
 1213 CCTCTTTGATTTTGGCTGGATTTATTTATTTATTTATTTTCTGAAATGTTGGGTAAAC 1272
 Db -----TAAAGATATTTTCTGCTGTTCTAGATGATGATTTCC 2146
 1273 ATCTTTAGATGAAGTTTCTCTAGCCTCTTT--AGTCTGCAATTTGAAGATAGATA 1330
 Db CTCTTTGTTGGAGTTGTCTCAATTTACTATCTCTTGAAGGGCTGGATTTGTGGAAGATA 2086
 1331 TTCTTTACATCTGATTTTATCTAGATGCTTTCTTTCTCCAACTATTCTGACAGAAAG 1390
 Db TTGTGTAATTTGTTTGTGATGGAATATCTGGTTTCTCCATCTATGTTAAATGAGAG 2026
 1391 TTTTCTTAAGTGCAGTGTGCGCTGACATCTGTAGTCTCTTTGGAGTCTGTAGCACATC 1450
 Db TTTTACTGGGTATAGTAGCTGCTCTGGCATTTGTGTCTCTTAGGACGTGTATGACATC 1966
 1451 TGTGACGGCTCTTACATTTGATTTTCTTATTTGGAAGTACGCTGTAATTTCTAATAC 1510
 Db TGCCAGGATCTTCTAGCTTTTCTGTTCTGTTT-GAGAAATCTGGTGTAAATTTCTTATAG 1907
 1511 ATCTGCCCTTATATGTTAAATTTGTTCTTTTTCCTCTGATCTTTTAAATATCTTTCTTTG 1570
 Db GACTGCGCTTAT---GTACTTGACCTTTTTCCTTACTCTGCTTTATTTCTTTCTTTG 1850

1571 TTCTATACCTTTTAGTGATTTGATTTATGACATGTGTGGGAGTTTCTTTTCCGTFCCAAT 1630
 Db TTTTATGCAATTTGGTGTCTTGGTTATTATGTGACAGGAAGAATTTCTTTTCTGGTCCAAT 1790
 1631 CTAATTTGGTGTGTGTGATGCTTCTGTACCTTGATAGGACATCTCTTCTCAAGCTTAGGA 1690
 Db CTAATTTGGAGTCTCTGAGACCTCTTGTATATATGATGGGTATCTATTTCTTTAGGTATGG 1730
 1691 AATTTTCTTTTGGTGTCTTCTGAAATATTTTCCCTGCTTTT-----GACCTGCC 1742
 Db AGTTTCTTCTATATATTTTATTTGAAGATATCTACTGCGCCCTTTAAGTTGGATCTTCT 1670
 1743 TTCTTCCCTCTCTCTATCTCTTGGTGTGTGTGATGATGCTGTCTGGCTTCTCTGATGTTT 1802
 Db TCTATACCTTATATCATAAAGTTTGGTCTTCTCATTTTGTCTGGATTTCTCTGATGTTT 1610
 1803 TATGCTGATTTATTTAGACTTAACATTTTCTTTTGACAAAGTATCCATTTCTTCTATC 1862
 Db TGGTTAGAGCTTTTACATTTTGCAATTTCTTTGACTGTGTGCAATGTTTCTTAGG 1550
 1863 TTGCTTTCACCTGCGGAGATTTCTTCTATCTTCTATCTTGTATCTGTGATGAGCTGTGTC 1922
 Db GCATCTTCTGACCTGGGATTTCTCTTTTATCTTCTTGTATCTTGTGTGATGCTTGA 1490
 1923 TCTGAGGTCTCTG-TTGGGTTCTTAAATTTTCTTATTTCCAGATTTCCCTCAGTTGGGTT 1981
 Db TCTATGATCTCAGATCTCTCTCTAGGTTTCTATCTCCAGGTTGTCTCCCTTTGTAT 1430
 1982 TTGTTTATTAATTTTATTTCCACTTTCCAGTCTCGGAAATGTTTACTCAATTTTCTCCCA 2041
 Db TTCTTTATTTGTTCTATTTTCCATTTTAGGCTCTGGAAGCTTTTATTTCAATCTCTCAC 1370
 2042 GATT-----TACATTTTCATAGTTT 2063
 1369 TGTTTATGTTATTTCTCTTTAAGACTTCTACCTGTGTGACCTGTCTCTCTGTTATTT 1310
 2064 CTTTAAATGATTTATTTCAATTTCTCTCAAGGACCTTTTATGAATTCATAAAATGTATGT 2123
 Db CATTAAAGGAGTATTTATGTTCTTAAAGTCTCTTATTTGGCATCATGATGATTAAT 1250
 2124 TAAGTCTCTCTCTGCTGCTTCAAGTATGTTGCAATTTCTCAGG 2166
 Db TAAATCCAAACCTTCTCTTTTGGGTGTGTGGAGTAGCCAGG 1207
 RESULT 19
 BZ124878/c
 LOCUS
 DEFINITION
 CH230-490E23.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
 CH230-490E23, genomic survey sequence.
 ACCESSION
 BZ124878
 VERSION
 BZ124878.1 GI:23765825
 KEYWORDS
 GSS.
 SOURCE
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 (bases 1 to 1151)
 AUTHORS
 Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
 Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
 Riggs,F., de Jong,P. and Fraser,C.M.
 TITLE
 Rat BAC End Sequences from Library CHORI-230 MboI segment
 JOURNAL
 Unpublished (1999)
 COMMENT
 Other GSSs: CH230-490E23.TJ
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library availability, please contact Pieter de Jong (pdejong@mail.choi.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or_ering_information.htm). BAC end page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 490 row: E column: 23
 Seq primer: 17
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1. 1151
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSNHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-490E23"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 2"
 /notes="Vector: pTARBAC1.3; Site 1: Mbol; Site 2: Mbol; CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by Pieter de Jong"

Query Match 17.5%; Score 399.299; DB 1; Length 1151;
 Best Local Similarity 63.7%; Pred. No. 9;
 Matches 775; Conservative 0; Mismatches 347; Indels 95; Gaps 12;

534 TTTCGTTTGGTGAATAGTCGTAAATATC-TCTAGGTCCACTGGTTTATGACATCA 592
 |||||
 1146 TTTCGTTTGGATAGATGCTTATAAATACGGTTAAGTCCAAATTTGGCTCATGACTTCT 1087
 |||||
 593 GTTAGCTCCAGCATTTCTCGTTTGGTTTGTGTGAGATGACCTAACTGTTGGAGAGAA 652
 |||||
 1086 CTTAGCTGCTACATCTCTGTTTAATTTCTGTT-CCAATGCTGTCCATGATGAGAG 1028
 |||||
 653 TGGGTAATCAAGTACGACCACTATCTGTGTGTGAGGT-CAATATGATATTTAGCTGTAG 711
 |||||
 1027 TGGGTGGGAATCTCCCACTATTATGTGTGAGGTGCAATGTGTGTTTGGACTTTCG 968
 |||||
 712 CTGTGCTGTTTATGAACTGGGTGACATTTGTTGGTGCATAGACATTAAGATTGC 771
 |||||
 967 TAAGTTTCTTTTACATATGATAGTGCCTTGTTATTTGGGGCAATAGATTTAGGATTGA 908
 |||||
 772 AATGCTCTCTTGTGGGA-TTTTCTTTGATGCCATGATAGTATTTCTTCCAAATCTCATCT 830
 |||||
 907 GAGTTTCACTTGTGGGATTTTCTTTGATGAATATGAAGTGCCTTCTCTTATCTTTT 848
 |||||
 831 GCTTAGTTTGGGTTTAACTGA-TTAGTCAGATATTAATGACTATCGGCTTGCCT 889
 |||||
 847 GATGACTTTTGTGTTGAAATTTGATTTTATTGTATATTAGAAATGGCTACTCCAGCTTGC 788
 |||||
 890 CTTAGGGCCATTGCTTAGAATATC-TTTTCCATCTTTTACTCTAAGGTGATGCTATC 948
 |||||
 787 CTTCTGACCATTTGTTGGAAATTTTCCAGCCTTTTCCATCTGAGGTAAATGCTGTC 728
 |||||
 949 CATG---GTAGTTGCTTTTGGATGCGACAGTAGATGGAATGTTGTTTCAATATCCA 1005
 |||||
 727 TTTGCTCTGAGGTGTGTTTCTCTAGGAGCAGATGCGAGGTCTCTGTTGGGTATCCA 668
 |||||
 1006 TTCTGTTTACCAGATCTTTTCTAGAGAAATTAAGATCATTTGATGATGTTGAGA 1065
 |||||
 667 GTTTGTTTAACTATGCTTTTTTATTTGGGGATTGAG-----GCCATTTGATTTGAGA 616
 |||||
 1066 ATTATCAATGACGAGTGTGTTGGATTTCTGTTATCTTGACCTTTGGAAGTGTGTGTG 1125
 |||||
 615 GATATTAAAGGAATAGTATTTATCTTCCCGTTATATTATTTGATTTGGATGAGGTTATG 556
 |||||
 1126 TGT 1185
 |||||
 555 TTTGTGTG-----
 |||||
 1186 TGT 1245
 |||||
 547 -----CTTCACTCTCTTTGTTTGTGCGCAAGATGATTAGTT 510
 |||||

QY 1246 ATTCAATATTTCTTGAATGTGGGTAAACATCTTTAGATTGAAGTTTTC--TCCTAGCCTT 1303
 Db 509 TCTTGTCTTCTTAGGTATAGCTTGCCTCCTTATGTGGCCTTTACCATTTCTTATCCT 450
 |||||
 QY 1304 CTTTAGGTCCTCATTGGAAGATAGATATTTTACATCTGATTTTATCTTAGAATGTCTT 1363
 Db 449 TTGTGGTCTGGATTTGTAGAAAGATATTGTGTAATTTGGTTTGTGATGATACTTT 390
 |||||
 QY 1364 TCTTTCTCCAACTATTTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCTGACATCT 1423
 Db 389 GGTTCCTCCGTCTATGTTAAATTGAGAGTTTTCAGAGATACAGTAACCTGGGCTGGCATTT 330
 |||||
 QY 1424 GTAGTCTCTTGGAGTCTGTACACATCTGTGACGGCCTTCTTACATTTTGTAGTTTCTAT 1483
 Db 329 GTGTTCTCTTAGAGTCTGTATGACATCATGTCACCAATCTTCCTGCTTCATAGTTTCT-G 271
 |||||
 QY 1484 TGGAAAGTCAGGTGTAAATCTAATACATCTGCCCTTTATATGTTAAATGGTCTTTTTTCC 1543
 Db 270 GCGAGAAGTCTGGTGTGATTTCTGATAGGTCTGCCCTTTATATGTTACTT-GACCTTTTCC 212
 |||||
 QY 1544 CTTGCACTTTTAAATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1603
 Db 211 CTTACTGCTTTTAAATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 152
 |||||
 QY 1604 TGTGGGGAGTTTCTTTTCCGGTCCAACTATTTTGGTGTGTTTCTGATGCTTCTTGACCTTG 1663
 Db 151 CGGAGAGTGTCTTCTTCTGCTCCAACTATTTGGAGTCTGTAGGCTTCTTGTATGCT 92
 |||||
 QY 1664 ATAGCATCTCTTCTCAAGGTAGGAAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1723
 Db 91 ATGGGTATCTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 32
 |||||
 QY 1724 TCCTGCTTTTGGACCTG 1740
 |||||
 Db 31 ACTGTCCTTTGAGCTG 15
 |||||

RESULT 20
 BZ135754
 LOCUS
 DEFINITION CH230-287J14.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
 CH230-287J14, genomic survey sequence.
 BZ135754
 VERSION BZ135754.1 GI:23776701
 KEYWORDS GSS.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 1065)
 AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
 Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, F., de Jong, P. and Fraser, C.M.
 TITLE Rat BAC End Sequences from Library CHORI-230 MboI segment
 JOURNAL Unpublished (1999)
 COMMENT Other GSSs: CH230-287J14.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.choi.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/or_ering_information.htm). BAC end
 page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 287 row: J column: 14
 Seq primer: SP6

Db	699	ATTGTGTTCTCTTAGGGTCTGTAAGCATCAGTCAGGATCTCTGGCCCTCATAGTTT	758
Qy	1480	CTATTGGAAAGTCAGGTGTAATTTCTTAATACATCTGCCCTTTATATGTTAAATGGTCTTTT	1539
Db	759	CT-GGCGAGAAGTCTGGTGTGATCTGATAGGTCTGCCCTTTATATGTTACTT-GACCTTT	816
Qy	1540	TTCCCTTGCACTTTTAAATATCTTTCTTTGTTCTATACATCTTTTAGTGATTCGATTATTAT	1599
Db	817	TTCCCTTACTGCTTTTAAATATCTTTCTTTTATTTTGTGGTGTGGTGTTCGACTATTAT	876
Qy	1600	GCACTGTGGGAGTCTTTTTCGGTCCCAATCTATTGTGGTGTTCGATGCTTCTGTAC	1659
Db	877	GTGACGGGAGGTCTTTCTTTCTGGTCCCAATCTATTGTGAGTTCGTAGGCTTCTGTAT	936
Qy	1660	CTTGATPAGGATCTCTTTCTCAAGGTAGGAAATTTTCTTTTGGTGTTCCTGAAAT	1719
Db	937	GCCTTAGGGTATCTCTTTTCTTTTAGGTAGGAAAGTTCCTCTATGATTTCTGTGAAGAC	996
Qy	1720	ATTTCCTGCTTTTGACCT-GCCTTCTTCCCTTCTCTATTCCTTGGTGTTCCTT	1773
Db	997	ATTTACTGCTCTTTGAGCTGGAGTCTTCACTCTCTCTATFACCTATTATACCTT	1051
RESULT 21			
LOCUS	AK040990	2372 bp mRNA linear HTC 19-SEP-2003	
DEFINITION	Mus musculus adult male aorta and vein cDNA, RIKEN full-length enriched library, clone:A530056124 product:unknown EST, full insert sequence.		
ACCESSION	AK040990		
VERSION	AK040990.1	GI:26088196	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	1	Carninci, P. and Hayashizaki, Y. Eukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	3	Carninci, P., Shibata, Y., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
AUTHORS	4	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	
TITLE	Genome Res. 10 (11), 1757-1771 (2000)		
JOURNAL	20530913		
MEDLINE	11076861		
PUBMED	11042159		
REFERENCE	5	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
AUTHORS	6	Functional annotation of a full-length mouse cDNA collection	
TITLE	Nature 409, 685-690 (2001)		
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE	7	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	
AUTHORS	8	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
TITLE			

JOURNAL
COMMENT

Unpublished (1999)
Other_GSSs: CH230-208116.TV3
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orering_information.htm). BAC end
page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Plate: 208 row: 1 column: 16
Seq primer: SP6
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..886
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-208116"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by
Pieter de Jong"

Query Match 15.9%; Score 363.399; DB 1; Length 886;
Best Local Similarity 70.4%; Pred. No. 16;
Matches 564; Conservative 0; Mismatches 221; Indels 16; Gaps 6;

QY 1239 ATTATTATTCATATTTCTTGATGTTGGGTAAACATCTTTAGATTGAAGTTTTCCTCTA 1298
DB 802 ATTACTTCTGCTTTTCTAGGCTGATGTTCTCTCTGTTGGAGTTTCCATCTA 743

QY 1299 --GCCTCTTTAGTCTGCAATTTGAAGATAGATATCTTTACATCTGATTTATCTAGA 1356
DB 742 TTATCATATATAGAGCTATATATATGAAAGATATTTGATAAAATTTGGTTTCTTGTA 683

QY 1357 ATGTCCTTTCTCCAACTATTCTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCCT 1416
DB 682 ATATCTTTGTTCTCCATCTATGTAATTTATAGTTTGTAGTATAGTCTGGCT 623

QY 1417 GACATCTGATCTCTTTGGAGTCTGTAGCACATCTGCGAGGCGCTTCTTACATTTTGA 1476
DB 622 GGCATGTGTCTCTTATAGGCTGTATGACATCTGTCCAGGATCTTCTGGCTTTCATAG 563

QY 1477 TTTCTATGGAAGTCCAGGTGAATCTTAATACATCTGCTTATATGTTTATGTTCT 1536
DB 562 TCGTAGT-GAGAGTCTGGTGAATCTTAATAGTCTGCTTTTATGTTACTT-GAGC 505

QY 1537 TTTTCCCTGTCATCTTTTAAATATCTTTCTTTGTTCTATACATTTTATGTTGATTTAT 1596
DB 504 TTTTCCCTTACTGCTTTTAAATATCTTTCTTTATTTGTGCAATTTGTTGTTGACTAT 445

QY 1597 TATGACATGTTGGAGTTCTTTTCCGGTCCAATCTATTTGTTGTTTGTATGTTCTTTG 1656
DB 444 TATGTTACGGAGAGGATTTCTTTCTGGTCCAATCTATTTGAAGTTCTGTAGGCTTCTTG 385

QY 1657 TACCTTGATAGGATCTCTTTCTCAAGTTTAGGAATTTTCTTTTGTGTTTCTTTGAA 1716
DB 384 TATGTTATGGCATCTCTTTCTTTAGTTAGGAAGTTTCTTCTATGATTTTGTGAA 325

QY 1717 AATATTTTCCCTGCTTTTGTACCTG-----CCTTCCTCCCTCTCTCTATTCCTTTG 1767
DB 324 GATATTTACTGGCTTTTGTAGCTGGAGTCTTCTACCTCTCTTATACCTATTTATCTTAG 265

QY 1768 GTTT--TTGCATAGTGTCTCTGGCTTCTCGATGTTTATGCTGGATTTATTTAGACTT 1825

DB 264 GTTGTATCTCATTTGAGTCTCTGATTCTCTGTATGTTTATAGCCAGTAGCTTTTCTCTTT 205
QY 1826 AACATTTTCTTTGACCAAGGTATCCATTTCTTATCTTCTTCTTCTTCTTCTTCTTCTTCT 1885
DB 204 TACATTAATCTTTGACCAATTTGTCGATGATTTCTATGGAATCTTCTGCTCTGATTTCT 145
QY 1886 CTCTTCTATCTCTTCTATTTCTGTCAGTGGCTTCTCTGAGGTTCC-TGTGGGTTCT 1944
DB 144 CTTCTCTATCTCTTGTATTTCTTGTGATGCTTGTATCTATGCTCTCTTCTCTTCT 85
QY 1945 TAAATTTTCTATCCAGATTTCTTCTAGTTGGTTTGTATTAATTTCTATTTCCAC 2004
DB 84 TTGGTTTCTATATCCAGATTTCTCTCTTGGCTTCTTCTTATTTGTTCTTATTTCCAT 25
QY 2005 TTTTCAGGTCCTGAAATGTTTT 2025
DB 24 TTTTAAATCTTTCACCTGTTT 4

RESULT 23
BH348678/c

LOCUS
DEFINITION
BH348678 846 bp DNA linear GSS 03-DEC-2001
CH230-42H13.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-42H13, genomic survey sequence.

ACCESSION
VERSION

BH348678
BH348678.1 GI:17279412

KEYWORDS
ORGANISM

Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 846)
Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
Shvartsbeyn, A., Gebregorgis, B., Overton, L., Russell, D., Chen, D.,
Riggs, P., de Jong, P. and Fraser, C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)

TITLE
JOURNAL

Other_GSSs: CH230-42H13.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

COMMENT

Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orering_information.htm). BAC end
page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Plate: 42 row: H column: 13
Seq primer: SP6
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..846
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-42H13"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by
Pieter de Jong"

Query Match 15.9%; Score 361.399; DB 1; Length 846;
Best Local Similarity 67.0%; Pred. No. 17;
Matches 567; Conservative 0; Mismatches 261; Indels 18; Gaps 4;

QY	1210	TCCTCTTTTGGATTTTGGCTGGGAATATTATTATTATTCATATTTCTTTGAATGTGGT	1269
Db	845	TTTCTTTCTTTGTTTGGCTGGGAATATTATTATTATTCATATTTCTTTGAATGTGGT	786
QY	1270	AAATCTTTAGATTTGAATTTTCTCTAGCTCTTT--TAGGTCTGCAATTTGAAGATAG	1327
Db	785	GCCTCTTTATTTGGATTTTACCAATTTATATCCCTTTGTAGTCTGATTTGTAGAAAG	726
QY	1328	ATATCTTTACATCTGATTTATTTATTTAGATTTCTTTCTTTCTTCCCACTATTGTGACAG	1387
Db	725	ATATTGTGAAATTTGGTTTGTATGGAATATCTTTGGTTCTTCCCACTATTGTATTTGA	666
QY	1388	AAATCTTTTCTAAGTGCAGTGTCTGGCTGCACATCTGTAGTCTTTCTGGAGTCTGTAGCAC	1447
Db	665	GAGTTTTCAGGATACAGTAACCTGGCTGCACATTTGTGTCTTTAGGCTGTATGAC	606
QY	1448	ATCTGTGAGGCTCTTACATTTTGTAGTTCTTATTTGGAAGTGTAGTGTATTTCTTAA	1507
Db	605	ATCAGTCCAGGATCTTCTGGCTCTCATAGTTTCT--GGCGAAAAGTGTGTGATTTCTGA	547
QY	1508	TACATCTGCTTTATATTTAATTTGTCTTTTCTTCTTCCCTTGCATCTTTTATATTTCTT	1567
Db	546	TAGTCTGCTTTATATTTTACTT--GACCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCT	488
QY	1568	TTCTTTCTATCTTTTATTTGATTTGATTTATTTATGCACTGTGGGAGTTCTTTTCGGTCC	1627
Db	487	TTATTTTGTGGTTTGTGTTTGTGACATTTATGTCAGGAGTCTTCTTTCTGCTCC	428
QY	1628	AACTATTTTGGTGTTTGTATGTTCTTCTTGTACCTTGTAGGATCTTCTTCAAGGTTA	1687
Db	427	AACTATTTTGGGTTCTGTAGGCTTCTTGTATCTGTATGCTGTATGCTTTTCTTTTAGGTTA	368
QY	1688	GGAAATTTTCTTTTGTGTTTCTTGAATAATTTTCCCTGCTTTTGACCTG-----	1740
Db	367	GGGAGTTTCTTCTATGATTTTGTGAGATATTTACTGGTCTTTGAGCTGGAGTCT	308
QY	1741	-----CCTTCTCCCTTCTCTATCTTACCTTATTTAGTCTTCTTCTTCTTCTTCTTCTT	1793
Db	307	TCACCTCTTCTTATACCTTATTTACCTTATTTAGTCTTCTTCTTCTTCTTCTTCTTCTT	248
QY	1794	TGATGTTTATGCTGGATTTATTTAGCTTAACTATTTCTTTTGTGACCAAGGATCCATT	1853
Db	247	TGTATGTTTGGATCAGTACTTTTCCGCTTTTACATTTATTTTGTGACAGTGTAGTCAATG	188
QY	1854	TCCTCTATCTGTCTTACCTGCTGAGATTTCTCTTCTATCTTCTTGTATCTGTCAGTG	1913
Db	187	ATTCTATGGAATCTTCTGCTCTGTAGATTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTT	128
QY	1914	AGGCTGCTCTGAGGTTCTGTTGGTCTTCTTAAATTTTCTTCTTCTTCTTCTTCTTCTT	1973
Db	127	AAGCTTGTATCTACAGCTCTCTTGTCTCTTCTTGGTTTCTTATCTTCTTCTTCTTCTT	68
QY	1974	TTTGGGTTTGTATTTATTTATTTCTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2033
Db	67	TGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	8
QY	2034	TCCTCC 2039	
Db	7	TTTCCC 2	

RESULT 24
BZ221947/c
LOCUS BZ221947 825 bp DNA linear GSS 11-OCT-2002
DEFINITION CH230-361C20.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-361C20, genomic survey sequence.
ACCESSION BZ221947
VERSION BZ221947.1 GI:23880305
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
Rattus.	
1 (bases 1 to 825)	
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de Jong,P. and Fraser,C.M. Rat BAC End Sequences from Library CHORI-230 MboI segment Unpublished (1999)	
Other_GSSs: CH230-361C20.TU	
Contact: Shaying Zhao	
Department of Eukaryotic Genomics	
The Institute for Genomic Research	
9712 Medical Center Dr., Rockville, MD 20850, USA	
Tel: 301 838 0200	
Fax: 301 838 0208	
Email: szhao@igr.org	
Clones are derived from the rat BAC library CHORI-230 (http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or ering information.htm). BAC end page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html Plate: 361 row: C column: 20 Seq primer: T7 Class: BAC ends.	
Location/Qualifiers	
1..825	
/organism="Rattus norvegicus"	
/mol_type="genomic DNA"	
/strain="BN/SsNhd/MCW"	
/db_xref="taxon:10116"	
/clone="CH230-361C20"	
/sex="Female"	
/cell_type="Brain"	
/clone_lib="CHORI-230 Segment 2"	
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; CHORI-230 Rat (BN/SsNhd/MCW) BAC library produced by Pieter de Jong"	
Query Match 15.7%; Score 358.399; DB 1; Length 825;	
Best Local Similarity 67.8%; Pred. No. 18;	
Matches 555; Conservative 0; Mismatches 246; Indels 18; Gaps 4;	
QY	1213 CCTCTTTGATTTTGGCGCTGGGAATATTATTATTATTCATATTTCTTGAATGTGGTAAAC 1272
Db	821 CTCTCTTGTGTTTGTGCCAAGACTATTAGTTCTTGTCTTTTCCAGGGTGTAGCTTGC 762
QY	1273 ATCTTTAGATTTGAAGTTTCTCTCTA--GCCTCTTTTAGGTCTGCATTTGAAGATAGATA 1330
Db	761 CTTTAAATGTTGGCTTTGCCATATATTACCTTTGTAGGCTGGATTTGTAGAAAGATA 702
QY	1331 TTCTTTACATCTGATTTTATCTTAGAATGTCTTTCTTCTCCAACTATTTGTGACAGAAAG 1390
Db	701 TTGTGTAATTTGGTTTGTCTATGGAATATCTTGGTTTCTCCATCTATGTTAAITGAGAG 642
QY	1391 TTTTCTTAAGTGCAGTGTCTGGCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATC 1450
Db	641 TTTTGTCTGGATACAGTAACCTGGCTGGCAITTTGTGTTCTCTTAGGGTCTGTATGACATC 582
QY	1451 TGTGAGGCGCTTCTTACATTTTGTAGTTTCTATTGGAAAAGTCAGGTGTAATCTTAATAC 1510
Db	581 TGTCCAGGATCTTCTGGCTTTTCATAGTCTCTGAT--GAGAAGTCTGGTGTATCTGTAG 523
QY	1511 ATCTGCTTTATATGTTAAATGTTGTTCTTTTCCCTTTGCATCTTTTAAATATCTTTCTTTG 1570
Db	522 GTCTGCTTTATATGTTTACTT--GACCTTTTCCCTTACTGCTTTTAAATATCTTTCTTTA 464
QY	1571 TTCTATACCTTTTGTAGTATTGATTATTATGCACTGTGGGAGTTTCTTTTCGGTCCCAAT 1630
Db	463 TTTTGTGCAATTTGGTGTGTTTGTACTATTATGTGACGGGAGGAGTTTCTTTTCTCGTCCAAT 404
QY	1631 CTATTGGTGTGTTTGTATGCTTCTTGTACCTTGTAGGCAATCTCTTTCTCAAGGTTAGGA 1690

CH230-394F19, genomic survey sequence.

BZ132018
BZ132018.1 GI:23772965

GSS.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 1037)

Zhao,S., Shetty,J., Shateman,S., Teegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999).
Other_GSss: CH230-394F19.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0300
Fax: 301 838 0308
Email: szhaoc@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).

Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Plate: 394 row: F column: 19
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..1037
/organism="Rattus norvegicus"
/mol_type="Genomic DNA"
/strain="BN/SNHsd/MCW"
/db_xref="taxon:101116"
/clone="CH230-394F19"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/notes=Vector: PIABAC1.3; Site:1: MboI; Site:2: MboI;
CHORI-230 Rat (BN/SNHsd/MCW) BAC library produced by
Pieter de Jong"

Query Match 15.6%; Score 354.899; DB 1; Length 1037;
Best Local Similarity 64.5%; Pred No. 15;
Matches 614; Conservative 0; Mismatches 301; Indels 37; Gaps 6;

Qy 1155 GTGTGTTGTCTCCCC 1214
Db 980 GTGATTATTCCTCCCTGTATTTCATATTTGGATGTGAGGTTATGTTTGTGTCTTCCA 921

Qy 1215 TCITTTGATTTTTGGCCTGGGAATATTATTATTCATATTTCTTGATCTGGGTACAT 1274
Db 920 TTCCTTTGTTGTGTGCAGATGATTAGTTCTTGTCTTCTTAGGGTATAGCTTGCGT 861

Qy 1275 CTTTAGATTGAAGTTTTTCTCCTAGCCTTCTT--TAGTCTGCATTTGAAGATAGATTT 1332
Db 860 CCITATCTTGGCCITTACCCTTTATTATCCITTCAGTCGTGGATTGTAGAAGATTT 801

Qy 1333 CTTTACATCTGATTTTATCTTTAGAAATGTCTTTCTTTCTCCCAACTATTGTGTGACAGAAAGTT 1392
Db 800 GTGTAAATTTGGTTTGTTCATGGAATATCTTGGTTCTCTCCATCATGTATTAATTGAGAGTT 741

Qy 1393 TTTCATAAGTCGAGTAGTCTGGCCGTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATCTG 1452
Db 740 TTGCAGATACAGTAACTGCGGCTGACATTTGTCTTCTTAGGGTCTGTATGACATCTG 681

Qy 1453 TGCAGGCCCTTCTTACATTTTGTAGTTCTTATTTGAAAAAGTCAGGTGAATTTCTAATACAT 1512
Db 680 TCCAGGATCTCTGGCTTTCATAGTTCTTCT--GGCGAAAGTCTGGTGTGATCTCATAGGT 622

[illegible]

Db 150 CATCTCTGTATTCTCGTGGTGAAGCTTGATCTACAGCTCCTGTCTCTTCTTTTGGTT 91

868 GTAGTGTGGATTGTAGAGATATTGTGTAATTTGGTTTGTCTCATGGAATATCTTGG 809
1366 TTTCCTCACTATTGTGACAGAAAGTTTCTAAGTCAGTAGTCTGGCCCTGACATCTGT 1425
808 TTTCCTCACTAAATGTTAATGAGAGTTTCTCGGATACAGTAATCTGGCTGGCAATTTGT 749
1426 AGTCTCTTGGAGTCTGAGACATCTGTGAGGCCCTCTTACATTTTGTGATTTCTATTG 1485
748 GTTCTCTTAGGGTCTGTATAACATCAGTCAGGATCTCTGGCCCTCATAGTTCT-GGC 690
1486 GAAAGTCAGGTGTAATCTAATACATCTGCTTTATATGTTAATGCTTTTTCCTTCCCT 1545
689 GAGAAGTCTGTGATCTGATAGGCTCTCCCTTATATGTTACTT-GACCTTTTCCCT 631
1546 TGCATCTTTAATATCTCTTCTTGTCTTACATTTTGTAGTATGATTTATATGCACTG 1605
630 TACTGCTTTTAATATCTCTTCTTATTTGTGCTTTTGGTGTGTTGTCATATGTCAGC 571
1606 TGGGAGTCTCTTCTCGGTCCTAATCTATTGTGTTTGTATGCTTCTTGTACCTTGAT 1665
570 GGAGGTGTTCTTCTGTCCTAATCTATTGAGTCTCTGTAGGCTTCTGTATGCTAT 511
1666 AGGATCTCTTCTCAAGTTAGAAATTTTCTTTTGTGTTTCTTGAATATTTTC 1725
510 GGGTATCTCTTTTATAGGTTAGGAAAGTTTCTCTCATGATTTTGTGAAGATATTTAC 451
1726 CCTGCTTTTGACCTG-----CCTTCTCCCTCTCTCTATTCCTTTGGTT 1771
450 TGGTCTTTGAGCTGGAGTCTTCAATCTCTCTATACCTATTATCTTGGTTGATCT 391
1772 TTGCATAGTCTCTGGCTTCTGGATGTTTATGCTGATATTTTGTAGACTTAACAT 1831
390 TCTCATGAGTCTGGATTTCTGTATGTTTGGACAGTAGCTTTTTCGGCTTACAT 331
1832 TTCTTTGACCAAGTATCCATTTCTCTATCTTCTCTCTCTGCTGAGATCTCTCTTC 1891
330 ATCTTTGACAGTTGAGTCAATGATTTCTATGGAATCTCTGCTCGAGATCTCTCTTC 271
1892 TATCTTGTATCTGTGAGGCTGTCTCTGAGGTTCTGTGGTTCTTAAATTT 1951
270 CATCTCTGTATCTGTGTTGAGGCTGTATCTACAGCTCTGTCTCTCTTTGGTT 211
1952 TTC-ATTTCCAGATTTCTTCAGTTTGGTTTGTATTAATCTATTTCCACTTCAG 2010
210 TTCTATATCCAGGTTGTTCATGTTCTCTCTGATTTGCTTATTTCCATTTTAA 151
2011 GTCTGAAATGTTTACTCTATTTCTCCAGTATTTACATTTTATAGGTTTCTTAAAT 2070
150 TTCTTCACTGTT-----TGATTTGTTTCTCGGATTTCTTTCAG 109
2071 GGATTTATTCATTTCTCTCAAGGCTTTTATGAAATTCATAAATGT 2119
108 GGATTTTGGATTCCTCTCTGTAGGCTTTTACATGTTTATTAATGTTT 60

RESULT 29
BZ108872/c 920 bp DNA linear GSS 11-OCT-2002
LOCUS CH230-240A15.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
DEFINITION CH230-240A15, genomic survey sequence.
ACCESSION BZ108872
VERSION BZ108872.1 GI:23749751
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 920)
Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K.,
Sivarsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment

JOURNAL
COMMENT

Unpublished (1999)
Other_GSSs: CH230-240A15.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
Plate: 240 row: A column: 15
Seq primer: SP6
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1..920
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SSHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-240A15"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by
Pieter de Jong"

Query Match 15.4%; Score 351.499; DB 1; Length 920;
Best Local Similarity 67.6%; Pred. No. 17;
Matches 561; Conservative 0; Mismatches 250; Indels 19; Gaps 5;
1213 CTTCTTTGATTTTGGCTGGAATATTATATATCATATTTCTTGAATGTGGTAAC 1272
844 CTTCTTTGTTGTGCCCCCAGTGAATTAGTTCTGCTTTTCTAAGGGGTAGCTGC 785
1273 ATCTTTAGATTCAGTTTCTCCAGCTTCTT--TAGGCTGCTATTTGAGATAGATA 1330
784 CTCCTTATGTTGGCTTTACCAATTTATATCTTTGTAGTGTGATTTTAGGAAGATA 725
1331 TTCTTTACATCTGATTTTATCTTAGAATGTCTTTCTTCCAACTATTGTGACAGAAAG 1390
724 TTATGTAATTTGGTTTGTGATGGAATATCTTGTCTCCATCTATGTTAATTGAGAG 665
1391 TTTTCTAAGTCAGTAGTCTGCGCTGACATCTGTAGTCTCTTGGAGTCTGTAGCACATC 1450
664 TTTTCTGGATACAGTAACCTGGCTGGCAATTTGTGTTCTCTTAGGGTCTGTATACATC 605
1451 TGTGAGGCCCTCTTACATTTTGTGTTTCTATTGAAAAGTCAGGTGTAATTTCTAAATAC 1510
604 TGTCCAGGATCTTGGCTTTCATATCTCCGGT-GAGAAATCTGTTGATTTCTGATAG 546
1511 ATCTGCTTTATATGTTAATTTGCTTTTTCCTTGTGATCTTTTAAATATTTCTTTG 1570
545 GTCTGCTTTATATGTTACTT-GACCAATTTCCGTTACTGCTTTTAAATATTTCTTCT 487
1571 TTCTATATCTTTTAGTATTTGATTTATATGCACTGTGGGAGTTTCTTTCCGGTCCAAAT 1630
486 TTTTGTGATTTGGTTTGTGTTTGTACTATTATGACAGGAGGAGTTTCTTTTCGGCCCAAT 427
1631 CTATTGTTGTTTGTATGCTTCTTGTACCTTGATGAGCATCTCTTCTCAAGTTAGGA 1690
426 ATATTGGAGTTCTCTATGCTTCTTGTATGTTATATGGCATCTCTTTCTTAGTTAGG 367
1691 AATTTTCTTTTGGTTTCTTGAATAATTTTCCCTGCTTTTACCTG----- 1740
366 AAGTTTCTCTATGATTTGTTGAAGATATTTACTGCTCTTTTACTGCTGGAGGCTTCA 307
1741 ----CCTTCTCCCTCTCTCTATTTCTTGTGTTTGTGATAGTCTCTGCTGCTCTG 1796

Db 306 CTCCTCTCTACACCTGTTATCTCTAGGTTTCATCTCTCTCATTGAGTCTCGGATTCCTGT 247
 Qy 1797 ATGTTTATGCTGGATATTTAGACTTAACATTTCTTTGACCAAGGTATCAATTTCT 1856
 Db 246 ATGTTTGGACCAAGTATTTTCCATTTTCCATTTCTTTGACTGTGTGCAATGATT 187
 Qy 1857 TCTATCTGCTTCACTGCGCTGAGATCTCTCTTATCTCTCTGATATCTGTCAGTGAGG 1916
 Db 186 TCTATGAACTCTGCTGCTGAGATCTCTCTTATCTCTCTGATATCTCTTTGATG 127
 Qy 1917 CTGCTCTGAGGTTCC-TGTTGGTCTTAAATTTTTCATTTTCCAGATTTCTTCAGTT 1975
 Db 126 CTGCTATCTGCTGCT 67
 Qy 1976 TGGTTTGTGTTTATTAATCTCTATTTCCACTTTTCAAGTCTCTGAAATGTTT 2025
 Db 66 TGTCTCTCTTATGCT 17

RESULT 30
 BZ109050/c
 LOCUS
 DEFINITION CH230-218C4.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
 CH230-218C4, genomic survey sequence.

ACCESSION BZ109050
 VERSION BZ109050.1 GI:23749929
 KEYWORDS GSS.

SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 823)
 AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
 Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
 Riggs,P. de Jong,P. and Fraser,C.M.
 Rat BAC End Sequences from Library CHORI-230 EcoRI segment
 Unpublished (1999)

TITLE Contact: Shaying Zhao
 JOURNAL Department of Eukaryotic Genomics
 COMMENT The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/orering_information.htm). BAC end
 page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
 Plate: 218 row: C column: 4
 Seq primer: SP6
 Class: BAC ends.

FEATURES Location/Qualifiers
 1..823
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SENHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-218C4"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 1"
 /note="vector: pTARAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by
 Pieter de Jong"

Query Match 15.4%; Score 351.099; DB 1; Length 823;
 Best Local Similarity 67.3%; Pred. No. 19;
 Matches 561; Conservative 0; Mismatches 239; Indels 33; Gaps 5;

Qy 1301 CTTCTTTAGGCTGCTGATTTGAAGATAGATATTTCTTTACATCTGATTTTATCTTGAATGT 1360
 Db 819 CTTTGTAGTCTGATTTGTAGAAAGATAATTGTAAATTTGGTTTGTTCATGGAATAT 760
 Qy 1361 CTTCTTTCTCCAACTATTTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCTGACA 1420
 Db 759 CTTGTTTTCTCCATCAATTTAAATTTGAGAGTTTGTCTGATACAGTAACTGGGCTGGCA 700
 Qy 1421 TCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGGCTTCTTACATTTTTCAGTTTC 1480
 Db 699 TTTGTGTTCTCTTAGGGTCTGTATGACATCAGTCCAGGATATTTGGGCTTCATAGTTTC 640
 Qy 1481 TATTGGAAAGTCAAGTGTAAATTTCTAATACATCTGCGCTTTATATGTTAAATTTGGTCTTTT 1540
 Db 639 T-GGCGAGAAGTCTGCTGATTTCTGTAGGTCTCCCTTTATATGTTACTT-GACCTTTT 582
 Qy 1541 TCCCTTGCATCTTTTAAATATTTCTTTTGTGTTCTATATCTTTTAGTGAATTTGATTATATG 1600
 Db 581 TCCCTTACTGCTTTTAAATATTTCTTTTATTTTGTGCGTTTGGTGTTCGACAAATATG 522
 Qy 1601 CACTGTGGGAGTTTCTTTTCCGTCCTTCTTCCCTTCTCTTATTTCC-----TTTG 1767
 Db 401 TTTACTGTTCTTTGAGCTGGAGTCTTCACTCTCTTCTATACCTATTATCTTAGGTTG 342
 Qy 1768 GTTTTGTGATAGTGTCTCTGCTTCCGGATGTTTATGCTTGGATTTATGCTTGGATTTATAGACITAA 1827
 Db 341 ATCTTCTCATTTGAGTCTCTGGATTTCTGTATGTTTGTGACCAAGTACGCTTTTCCACITTA 282
 Qy 1828 CATTTTCTTTTGACCAAGGTATCCATTTCTTCTATCTTCTTCTCACTGCTGCTGAGATTTCTCT 1887
 Db 281 CATTATCTTTGACAGTTGAGTCAATGATTTCTATGGAATCTTCTGCTCTGAGATTTCTCT 222
 Qy 1888 CTTCTATCTCTGTATTTCTGTCAGTGGCTGTCTGAGGTTCTCTGAGGTTCTCTTGGTCTTTAA 1947
 Db 221 CTTCCATCTCTGTATTTCTGTTGGTGAAGTTGTATCTACAGCTCTTGTCTCTCTCTTT 162
 Qy 1948 TTTTTC-ATTTCCAGATTTCTTCCAGTTTGGTTTGTGTTTATTAATTTTATTTCCACTT 2006
 Db 161 GGTTTTCTATATCCAGGGTGTGTTTCCATGTTCTTTCTGATGCTTCTATTTTCCATTT 102
 Qy 2007 TCAGTCTCTGAAATGTTTACTCAATTTCTCCAGTATTTACATTTTCAAGTTTTCATAGGTTTCTT 2066
 Db 101 TTAATTTCTTCAACTGTTT-----GATTGTGTTTCTCTTGGAAATTTCT 59
 Qy 2067 TATGATTTATTTCAATTTTCTTCTTCAAGGACCTTTTATGAATTCATAAAATGT 2119
 Db 58 TCAGGGAATTTTGTGCTCTCTCTCTATGAGGCTTCTACTGTTTATTTATGTTT 6

RESULT 31
 BZ184947/c
 LOCUS
 DEFINITION CH230-329122.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
 CH230-329122, genomic survey sequence.
 ACCESSION BZ184947
 VERSION BZ184947.1 GI:23835048
 KEYWORDS GSS.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 816)
 AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,

Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, F., de Jong, P. and Fraser, C.M.
 Rat BAC End Sequences from Library CHORI-230 MboI segment
 Unpublished (1999)
 Other GSSs: CH230-329122.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/orering_information.htm). BAC end
 page: http://www.tigr.org/tadb/bac_ends/rat/bac_end_intro.html
 Plate: 329 row: 1 column: 22
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source
 1..816
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSNHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-329122"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 2"
 /notes="vector: PTARAC1.3; Site_1: MboI; Site_2: MboI;
 CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
 Pieter de Jong"

Query Match 15.4%; Score 349.899; DB 1; Length 816;
 Best Local Similarity 69.7%; Pred. No. 19;
 Matches 525; Conservative 0; Mismatches 211; Indels 17; Gaps 4;

QY 1297 TAGCCTTTAGTCTGCAATTTGAAGATAGATATCTTTACATCTGATTTATCTTAGA 1356
 |||||
 DB 782 TAATCCCTTTAGTCTGGAGTTGGAGAAAGATATTGTACATTTGGTTTGTCTATGGA 723
 |||||
 QY 1357 ATGCTTTCTTCCCAACTATGTGACAGAAAGTTTTCTAAGTCGACGAGTCTGGCCT 1416
 |||||
 DB 722 ATTTCTTGGTTTTCCCATCTATGTTAATGGAGATTTTGACGATACAGTAACCTGGCT 663
 |||||
 QY 1417 GACATCTGACTCTCTTGGAGTCTGTAGCACATCTGTGACGGCTCTTTACATTTTGAG 1476
 |||||
 DB 662 GGCATTTGTGTTCTCTTAGGTCGTGTATGACATCTGTCCAGGATCTCTGGCTTTTCATAG 603
 |||||
 QY 1477 TTTCTATTGGAAGTCAAGGTGTAATCTTAATACATCTGCCTTTATATGTAATGTGCT 1536
 |||||
 DB 602 TTTCTGCT-GAGAACTCTGGTGAATCTTATAGTCTGCCTTTTATGTTACTTGATC- 545
 |||||
 QY 1537 TTTTCTCCCTGCATCTTTTAAATATCTTTCTTGTCTTACATTTAGTATTTGATTAT 1596
 |||||
 DB 544 TTTTCTCCCTACTACTTTTAAATATCTTTCTTGTGTCATTTGGATTTTGTGACTAT 485
 |||||
 QY 1597 TATGCACTGTGGGAGTTCTTTTCGGTCCAATCTATTATGGTGTGTTGATCTCTTG 1656
 |||||
 DB 484 TATGTGACAGAGGAGTTCTCTTCTGGTCCAATCTATTGGAGTTCTGTAGGCTCTTG 425
 |||||
 QY 1657 TACCTTGATAGGATCTCTTCTCAAGTTAGGAATTTTCTTTTGTGTTTCTTGAA 1716
 |||||
 DB 424 TATGTTTATGGGATCTCTTTCTTTAGGTTAGGAAGTTTCTTCTACAAATTTTGTGAA 365
 |||||
 QY 1717 AATATTTTCCCTGCTTTTGACCTG-----CCTTCTTCCCTTCTCTATTC 1762
 |||||
 DB 364 GATATTTTCTGCTTTGAGCTGGAGTCTTCACTCTCTCTATACAGTATATCTTAG 305
 |||||
 QY 1763 CTTTGGTTTTGCAATAGTGTCTCTGGCTTCTGAGATGTTTATGCTCGGATATTTTGA 1822
 |||||

DB 304 GTTTGACCCCTCATTTGTCTCTGGAATTTCTGTATGTTTGGGCCACATCTTTCTG 245
 |||||
 QY 1823 CTTACATTTCTTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCTACCTGCTGAGAT 1882
 |||||
 DB 244 TTTTACATTAATCTTCATCAGTTGTGTCGATGATTTCATGCAATCTTCGGTCTCTGAGAT 185
 |||||
 QY 1883 TCTCTCTTCTATCTCTTGTATTTCTGTCTAGTGAGGCTTCTCTCTGAGTTCC-TGTGGGT 1941
 |||||
 DB 184 TCTCTCTTCTATCTCTTGTATTTCTGTGTTGTTGATCTGATCTGCGGCTCTCTGTCTCT 125
 |||||
 QY 1942 TCTTAATTTTTCATTTCCAGATTTCTTTCAGTTGGGTTTGTGTTTATTAATCTTATTTT 2001
 |||||
 DB 124 CTTTGGTTTTCTATATACCGGGTTTTCCTTTTGTCTTCTTCTTATTTACTTCTATTTT 55
 |||||
 QY 2002 CACTTTCAGGTCCTGAAATGTTTACTCATTTT 2034
 |||||
 DB 64 CATTTTCAATTCCTTCACCTGTTTGATTTGTT 32
 |||||

RESULT 32
 BZ097017/c
 LOCUS
 DEFINITION
 CH230-22306.TJB CHORI-230 Segment 1 Rattus norvegicus genomic clone
 CH230-22306, genomic survey sequence.

ACCESSION
 BZ097017
 VERSION
 BZ097017.1
 GI:23737901
 KEYWORDS
 GSS.
 SOURCE
 Rattus norvegicus (Norway rat)

ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS

TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 910)
 Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
 Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
 Riggs, F., de Jong, P. and Fraser, C.M.
 Rat BAC End Sequences from Library CHORI-230 EcoRI segment
 Unpublished (1999)
 Other GSSs: CH230-22306.TJB
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (<http://www.chori.org/bacpac/rat230.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/orering_information.htm). BAC end
 page: http://www.tigr.org/tadb/bac_ends/rat/bac_end_intro.html
 Plate: 223 row: 0 column: 6
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source

1..910
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SSNHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-22306"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 1"
 /note="vector: PTARAC2.1; Site_1: EcoRI; Site_2: EcoRI;
 CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
 Pieter de Jong"

Query Match 15.3%; Score 349.699; DB 1; Length 910;
 Best Local Similarity 67.3%; Pred. No. 17;
 Matches 561; Conservative 0; Mismatches 253; Indels 19; Gaps 5;

QY 1210 TCCCTCTTTTGGCTGGAATATTATATTCATATTTTCTTGAATGGGT 1269
 Db 846 TTATTCCTTTTGTGCCCCAAGAGTAGTTCTCTTCTCTAGGAGTAGCT 787
 QY 1270 AACATCTTTAGAGTAAATTTTC -TCTAGGCTTCTTTAGGCTGCAATTTGAAGATAG 1327
 Db 786 GCCTCCCTTATGTTGGCTTACCATTTATATCCCTTTGAGTCTGGAATTTGAGAAG 727
 QY 1328 ATATCTTTACATCTGATTTTACCTTTAGAGTCTTTCTTCTCCCACTATTCTGACAGA 1387
 Db 726 ATATGTTGAATTTGTTTGTTCATGGAATATCTGGTTTCTCCCACTATTGTAATGA 667
 QY 1388 AAGTTTTTAAAGTSCAGTAGTCTGGCTGACATCTGAGTCTCTTGGAGTCTGTAGCAC 1447
 Db 666 GAGTTTTTGTGATACAGTAACTGGCTGGCAATTTGTTCTCTTAGGCTGTATGAC 607
 QY 1448 ATCTGTCGAGGCTCTCTTACATTTTGAATTTCTATTTGGAAGTCAAGTGTAAATCTAA 1507
 Db 606 ATCAGTCCAGGATCTCTGCTCTCATAGTTCT -GGCGAAGTCTGGTGTGATCTGA 548
 QY 1508 TACATCTGCCCTTATATGTTAAATGGTCTTTTTCCTTGCATCTTTTAAATATCTTTCT 1567
 Db 547 TAGTCTGCTTATATGTTACTT -GACCTTTTCCCTTACTGCTTTTAAATATCTTTCT 489
 QY 1568 TTGTTCTATATCTTTTATGATTTGATTTATGACATCTGGGAGTCTTTTCCGGTCC 1627
 Db 488 TTAATTTCTGCTTGTGTTTGTGACAAATATGTGACGGGAGTGTCTTTTCTGCTCC 429
 QY 1628 AATCTATTTGTTGTTTGTATGCTTCTGTTACCTTGATAGCATCTTTCTCAAGSTTA 1687
 Db 428 AATCTATTTGAGTCTGTAGGCTCTGTATGTTATGAGTATCTTTTCTTTTATGTTA 369
 QY 1688 GGAAATTTTCTTTTGTGTTTCTTGAAATATTTTCCCTGCTTTTGACCTG ------ 1740
 Db 368 GGGAGTTTCTTCTATGATTTTGTGGAAGATTTTACTGCTCTTTGGGCTGGAGTCT 309
 QY 1741 -----CCTTCTCCCTTCTATCTTCTTTGGTTTGTGATGCTCTGGCTTCC 1793
 Db 308 TCACCTCTCTATACCTATTTCTTGTAGTTTGAICTTCTCATGAGTCTGGATTTCC 249
 QY 1794 TGGATGTTTATGCTCGGATTTATTTAGACTTAACTTTCTTTTGACCAAGGTATCCAT 1853
 Db 248 TGTATGTTTGGACAGTAGCTTTTCCGCTTTACATTAATCTTTGACGTGAGTCAATG 189
 QY 1854 TCTTCTATCTGCTTCACTGCTGAGATCTCTTCTTCTATCTCTTGTATCTGTGAGT 1913
 Db 188 ATTTCTATGAAATCTTCTGCTCTGAGATCTCTCTTCCATCTCTGTATCTGTGTTGGT 129
 QY 1914 AGGCTTCTCTGAGTCTCTGTTGGTCTTAAATTTTTC -ATTTCCAGATTTCTTCA 1972
 Db 128 AAGCTTGATCTACAGTCTCTGCTCTTCTTTTGGTTTCTATATCTTATCTGAGTGTCT 69
 QY 1973 GTTTGGGTTTGTATTAATTTCAATTTCCACTTTCCAGGCTGCTGAAATGTTT 2025
 Db 68 ATGAGTCTTCTTGTATGTTCTTATTTCCATTTTAAATCTTCAACTGTT 16

RESULT 33
 BZ244766/c
 LOCUS CH230-419B3.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
 DEFINITION CH230-419B3, genomic survey sequence.
 ACCESSION BZ244766
 VERSION BZ244766.1
 KEYWORDS GI:23905030
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1061)
 AUTHORS Zhao,S., Shetty,J., Shatsman,S., Teegave,G., Geer,K.,
 Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,

Riggs,F., de Jong,P. and Fraser,C.M.
 Rat BAC End Sequences from Library CHORI-230 MboI segment
 Unpublished (1999)
 Other_GSSs: CH230-419B3.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 938 0200
 Fax: 301 938 0208
 Email: szhao@tigr.org
 Clones are derived from the rat BAC library CHORI-230
 (http://www.chori.org/bacpac/rat230.html). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/or ering information.html). BAC end
 page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
 Plate: 419 row: B column: 3
 Seq primer: 77
 Class: BAC ends.

FEATURES
 Source
 1. 1061
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /strain="BN/SNHsd/MCW"
 /db_xref="taxon:10116"
 /clone="CH230-419B3"
 /sex="Female"
 /cell_type="Brain"
 /clone_lib="CHORI-230 Segment 2"
 /note="Vector: pTARAC1.3; Site 1: MboI; Site 2: MboI;
 CHORI-230 Rat (BN/SNHsd/MCW) BAC library produced by
 Pieter de Jong"

Query Match 15.3%; Score 349.499; DB 1; Length 1061;
 Best Local Similarity 64.7%; Pred. No. 15;
 Matches 618; Conservative 0; Mismatches 300; Indels 37; Gaps 7;

QY 1195 TGTGCTCTGCTCTCCCTCTTTTGTATTTTGGCTGGAATTTATTTATTCATATT 1254
 Db 960 TATGTTTGTGCTCTTCTCTTCTTTTGTGCTCAAG-ACGATTAGTTCTTGCTTC 902
 QY 1255 TCTGTATGTTGGGTACATCTTTAGATTGAAGTTTCTCTCTAGCTCTTT--TAGTTC 1312
 Db 901 TTTAGGTTAGCTTGCTCTTATGTTGGCTTTACCTTTATTAATCTTTGTAGTC 842
 QY 1313 TGCATTTGAAGATAGATATCTTTACATCTGATTTTATCTTAGAATGCTTTCTTTCTCC 1372
 Db 841 TGGATTTCTAGAAAGATATGTTGTAACCTTGGTTTGTCAATGAATATCTTGGTTCTCC 782
 QY 1373 AACTATTGTACAGAAAGTTTCTTAAGTGCAGTAGTCTGGCTGACATCTGTAGTCTCT 1432
 Db 781 ATCTATGTTAATGAGAGTTTTCAGGATACAGTAACCTGGGCTGGCATTTGTCTCT 722
 QY 1433 TGGAGTCTGTAGCACATCTGTGACGGCTTCTTACATTTTGAATTTTCTATTTGAAAGT 1492
 Db 721 TAGGCTCTGATGACATCTGTCCAGGATCTTCTGGCTTCCACATTTCT-GGCGAAAGT 663
 QY 1493 CAGGTGAATTTCTAATACATCTGCTTTTATATGTTAATTTGCTCTTTTCCCTGCTATCT 1552
 Db 662 CTGGTGTGATTTCTGATAGTCTGCTTTTATATGTTACTT-GACCTTTTCCCTTACTGCT 604
 QY 1553 TTTAATATCTTCTTTTCTTATCTTATCTTATGTTGCTTTTGTGCTTTTATGTCAGCGGAG 1612
 Db 603 TTTAATATCTTCTTTTATTTTGTGCTTTTGTGCTTTTATGTCAGCGGAGTG 544
 QY 1613 TTTCTTTTCCGTTCCAATCTAATTTGGTGTGTTGATGCTTCTTGTACCTGTGATAGGATC 1672
 Db 543 TTTCTTTTCTGCTCCAATCTAATTTGGAGTCTGTAGGCTTCTTGTATGCTATGGTATC 484
 QY 1673 TCTTCTCAAGGTTAGGAATTTTCTTTTGTGTTTCTTGTGAAATATTTTCCCTGCTT 1732
 Db 483 TCTTTTGTAGGAGTCTTCTTCTATGATTTTGTGAAAGATATTACTTGGTCT 424


```

QY 1733 TTGACCTGC-----CTTCTCCCTTCCTCTATTCTCTTGGTTTGGCATAG 1779
Db 423 TTGAGCTGGGAGCTTCACTCTCTCTATACCTATTATCTTAGGTTTGATCTTACATG 364
QY 1780 TGTCTCTGGCTTCTCGTGAATTTATGCTCGGATTTATTTAGACTTAAACATTTCTTTGA 1839
Db 363 AGTCTGGGATTTCTGTATGTTTGGACCACTAGTCTTTTACCGCTTTACATTATCTTTGA 304
QY 1840 CCAAGGATCAATTTCTTCTATCTTCTCTACGCTGAGATCTCTCTCTATCTCTT 1899
Db 303 CAGTTGAGTCAATTTCTTATGGAATTTCTGCTCTGAGATCTCTCTTCCATCTCTT 244
QY 1900 GTATTTCTGTCAGTGAGCTTCTCTGAGGTTCTCTGTTGGTTCTTAAATTTTTC-ATT 1958
Db 243 GTATTTCTGTTGGTGAAGCTGATGACAGCTCTCTGTCCTCTTTTGGTTTCTATAT 184
QY 1959 CCGAGATTTCTCGATTTGGGTTTGTATTAATTAATTAATTAATTAATTAATTAATTA 2018
Db 183 CCAAGGTTGTTTCCATGTTTCTTCTGATGCTTCTATTTCCATTTTAAATTCCTTCA 124
QY 2019 ATGTTTACTCAATTTCTCTCCAGTATTACATTTTTCATAGGTTTCTTTAAATGGAATTTAT 2078
Db 123 ACTGTT-----TGATTTGTTTCTCGGAATTTCTTTCAGGGAATTTT 82
QY 2079 TCATTTCTCTTCAAGACCTTTTATGAATTCATAAAATGATGTTAAGTCTT 2133
Db 81 GCGATTCCTCTCTGATGCTTATCTTCTCTTAAGGAGTTTCTTTATGTCCTT 27

RESULT 34
BZ272430 830 bp DNA linear GSS 15-OCT-2002
LOCUS CH230-263N19.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone
DEFINITION CH230-263N19, genomic survey sequence.
ACCESSION BZ272430
VERSION BZ272430.1 GI:23991968
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 830)
Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 MboI segment
Unpublished (1999)
Other_GSSs: CH230-263N19.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.choi.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end
page: http://www.tigr.org/tadb/bac_ends/rat/bac_end_intro.html
Plate: 263 row: N column: 19
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. 830
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SENHsd/MCW"
/db_xref="taxon:10116"
/clones="CH230-263N19"

```

FEATURES

source

```

BZ103057 878 bp DNA linear GSS 10-OCT-2002
LOCUS CH230-238N16.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
DEFINITION CH230-238N16, genomic survey sequence.
ACCESSION BZ103057
VERSION BZ103057.1 GI:23743936
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)

```

RESULT 35

BZ103057/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

```

/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 2"
/note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by
Pieter de Jong"

```

```

Query Match 15.3%; Score 349.099; DB 1; Length 830;
Best Local Similarity 70.6%; Pred. No. 19;
Matches 530; Conservative 0; Mismatches 204; Indels 17; Gaps 5;

QY 1301 CTTCTTTAGGCTGCAATTTGAAGATAGATATCTTTACATCTGATTTTATCTTGAATGT 1360
Db 45 CTTTGTAGGGCTGCAATTTGTAGAAAGATATTGTGTAAATTTGGTTTGTGATGGAATAT 104
QY 1361 CTTTCTTCTCAACTATTGTGACAGAAAGTTTTTCTAAGTCAGTAGTCTCTGGCTGACA 1420
Db 105 CTTGCTTCTCATCTATGTTAAATGAGAGTTTTGCTGGATACAGTAACCTGGGCTGGA 164
QY 1421 TCTGTAGTCTCTTGGAGTCTGTAGCACATCTGTGAGGCTTCTTACATTTAGTTTC 1480
Db 165 TTTGTGTTCTCTAGGCTCTGTATGAGATCTGTCCAGGATATCTGGCTTTTCATAGTCTC 224
QY 1481 TATTGAAAGCTCAGCTGTAATTTCTAATACATCTGCCCTTATATGTTAATGGTCTTTT 1540
Db 225 TGGT-GAGAGTCTGGTGTAAATCTGATAGTCTGCCCTCATATGTTACTT-GACCTTTT 282
QY 1541 TCCCTTGTCACTCTTTTAAATATTCTTTCTTTGTTCTATACCTTTTAGTGATTTATATG 1600
Db 283 TCCCTTACTGCTTTTAAATATTCTTTTGTGTCATTTGCTGTTTGAATATTATG 342
QY 1601 CACTGTGGGAGTTTCTTTTCGGTCCAAATCTATTGTTGTTTGTATGCTTCTGTACC 1660
Db 343 TGACCGGGAGGAGTTTCTTTTCTGTGTCCTAATCTATTGAGTCTGTAGACTTCTGTATG 402
QY 1661 TTGATAGGCACTCTCTTCTCAAGGTTAGGAAATTTTCTTTTGGTTTCTTGAATAA 1720
Db 403 TTTATGGGCACTCTTTTCTTAGGTTAGGAAATTTTCTTCTAUGATTTTGTGAGATA 462
QY 1721 TTTTCCCTGCTTTTGAACCTGC-----CTTCTCCCTTCTCTATTCCTTTG 1767
Db 463 TTTTCTGTGCTTTTGAAGCTGGAGTCTTCACTCTCTCTATACCTATATCCGTAGTTTG 522
QY 1768 GTTTTGTGATAGTCTCTGCTTCTGATGCTTTTATGCTGGATTTATTTAGACTTAA 1827
Db 523 ATCTTCTCATTTGCTGCTGGAATTTCTGTATGTTTGGACCAATAGCTTTTCCGTTTAA 582
QY 1828 CATTCTTTTGACCAAGGTATCCATTTCTCTATCTTCTCTCTCTGCTGCTGAGATTCCT 1887
Db 583 CATTATCTTTGACAGTTGTGTCGATGTTCTATGGAATCTTCTGCTCCGGAGATTCCT 642
QY 1888 CTTCTATCTCTGTATTTCTGTCAGTGAGGCTGCTCTGAGTTCC-TGTTGGGTTCTTA 1946
Db 643 CTTCTATCTCTTCTATTTCTGTTGGAGATGCTGTATCTGCGGCTCCTTTGCTCTTCCCTT 702
QY 1947 ATTTTTCATTTCCAGATTTCTCTCAGTTTGGTTTGTGTTTATTAATTTCCACTT 2006
Db 703 GGTTTCTATATCCAGAGGTGTCCTCCCTTGAGATTTCTTTATGCTTCTATTTCCATT 762
QY 2007 TCAGTCTC-TGAAATGTTTTTACTCATTTTC 2036
Db 763 TTAATTCCTTGACCTGTTTGAAGTTTTC 793

```

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 878)
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
Riggs, F., de Jong, P., and Fraser, C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL Unpublished (1999)
COMMENT Other GSSs: CH230-238N16.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@email.cho.org).
Clones may be purchased from BACPAC Resources
(<http://www.chori.org/bacpac/orering-information.htm>). BAC end
page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Plate: 238 row: N column: 16
Seq primer: T7
Class: BAC ends.

FEATURES

source Location/Qualifiers
1..878
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SnHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-238N16"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/notes="Vector: pPARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SnHsd/MCW) BAC library produced by
Pieter de Jong"

Query Match 15.3%; Score 348.299; DB 1; Length 878;
Best Local Similarity 67.6%; Pred. No. 18;
Matches 556; Conservative 0; Mismatches 247; Indels 19; Gaps 5;

QY 1215 TCCTTTGATTTGGCGTGGAAATATTATATATCATATTTCTTGAATGTGGTAACAT 1274
DB 821 TCCTCTGTTGGTGGCCCAAAATGATTAATTCCTGTTCTTTCTAGGGTATAACTTGCCT 762
QY 1275 CTTTAGATTCAGATTTTCTCCAGCTTCTT--TAGGTGCTGATTTGAAGATAGATATT 1332
DB 761 CTTATGTTGGGCTTTACCATATATTAATCTCCAGCGGTAGATTTGTAGAAGATATT 702
QY 1333 CTTTACATCTGATTTATCTTAGAATGCTTTCTTCTCCAACTATTGTGACAGAAAGTT 1392
DB 701 GTGTAAATTTGGTTTGTGATGAATATCTTGGTTCTCCATCTATGTTAACTGAGATT 642
QY 1393 TTCTAAGTCAGTAGTCTGGCTGACATCTGTAGTCTCTTGGAGTCTGTAGCAGATCTG 1452
DB 641 TTGAGGATACAGTAACCTGGGTGGCAATTTGTCTCTCTAGGGTCTGTATGACATCAG 582
QY 1453 TGACAGGCCCTTCTACATTTTGTAGTTTCTATTGGAAGATGAGTGTGTTCTTAATACAT 1512
DB 581 TCAGGATCTCTGGCTTTCATAGTTCT--GGAGAGAGTCTGTGTGATCTGTAGT 523
QY 1513 CTGCGTTTATATGTTAATTCGGTCTTTTTCCTTCGATCTTTTAAATCTTTCTTTGTT 1572
DB 522 CTGCGTTTATATGTTACTT--GACCTTTTCCCTTACTGCTTTTAAATCTTTCTTTTAT 464
QY 1573 CTATACCTTTTGTAGTTTGAATTAATGCACTGTGGGAGTTTCTTTTCCGGTCAATCT 1632
DB 463 TTGTGCAATTTGGTGTGTTTGTACTATTATGACGGGAGGTGTTCTTTCTGTGTCATCT 404

QY 1633 ATTTGGTGTGTTTGTATGCTTCTTGTACCTTGTATAGGCACTCTCTTCTCAAGTTAGGAAA 1692
DB 403 ATTTGGAGTCTGTAGGCTTCTTGTATGCCATATGGTAATCTCTTTTGTAGTTAGGAAA 344
QY 1693 TTTTCTTTTGTGTTTCTTGTAAATAATTTCCCTGCTTTTGACCTG----- 1740
DB 343 GTTTCTTCTATGATTTTGTGTGAAGATATTTACTGGTCTTGTAGCTGGGAGTCTCACT 284
QY 1741 --CCTTCTTCCCTTCCCTTATTCCTTTGGTGTGTTTGTAGTGTCTCTGGGCTCCTCGAT 1798
DB 283 CTCCTTCTATACCTATATCTTCTAGGTTTGTATCTTCTCAATGAGTCTCGATTTCTGTAT 224
QY 1799 GTTTTATGCTCGGATTAATTTTAGACTTAACATTTCTTTTGACCAAGGTATCCATTTCTTC 1858
DB 223 GTTTTGGACCAATAGCTTTTTCGCTCTACATTAATTTTGACAGTGTAGTCAATGATTTTC 164
QY 1859 TATCTTGTCTTCACTGCTGAGATCTCTCTCTATCTCTTGTATCTGTCTGAGTGAAGCT 1918
DB 163 TATTTAAATCTTCTGCTCCTGAGATCTCTCTCTATCTCTTGTATCTGTGTGTTGTTG 104
QY 1919 TGCTCTCAGGTTCTGTTGGTCTTAAATTTTTC-ATTTCCAGATTTCTTCTCAGTTTG 1977
DB 103 CGTATCTACAGCTCTTGTCTCTCTCTGTTTCTATATCCAGGGTGTGTTTCCAATGTG 44
QY 1978 GGTGTTGTTTATTAATCTATTTCCACTTTTCAGGTCCTGAAA 2019
DB 43 TTCTTTCTGTATGCTTCTTCTATTTCCATTTTAAATTTCTTCAA 2

RESULT 36

AK088420/c 2559 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN
DEFINITION full-length enriched library, clone:E430016D20 product:similar to
PORF1 [Mus musculus domesticus], full insert sequence.

ACCESSION AK088420
VERSION AK088420.1 GI:26104798
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL
REFERENCE
AUTHORS

Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 2559)

JOURNAL
REFERENCE
AUTHORS

1412 GGCTGACATCTGTAGTCTCTTTGGAGTCTGTGACACATCTGTGAGGCGCTTCTTACATT 1471
1315 CACCTGGCAATTTGTCTTCTTAGGGTATGTATGACATCTGCCAGGATCTTCTAGCTTT 1256
1472 TTGAGTTTCTATTGSAAGTCAGGTGAATTTCTAAATACATCTCCCTTTATATGTTAAT 1531
1255 CATGCTCTCGGT-GAGAAGTCTGTGTAAATCTGATAGGCTGCCCTTATATGTTACTT 1197
1532 GGTCTTTTCCCTTGGCATCTTTTAAATATCTTTCTTTGTCTATACATTTTGTAGTTG 1591
1196 -GACCTTTTCCCTTACCATTTAAATATCTTTCTTTGTCTAGTGCATTTGGTGTTCG 1138
1592 ATTATATGCACTGTGGGAGTCTTTTCCGGTCCATCTATTTGGTGTCTTCTATGCT 1651
1137 ATTATTAAGTGAATGGGGAATTTCTTTTCTGGTCCAGTCTATTTAGAGTTCTGTAGCT 1078
1652 TCTGTACCTTCTGATAGGCATCTCTTTCTCAAGGTGTAGAAATTTTCTTTTGTGTTTC 1711
1077 TCTGTATGTTTCATGGGCATCTCTTTCTTAGTGTAGGGAGTTTCTTCTATATTTG 1018
1712 TGAATAATTTTCCCTGCTTTTGCCTG-----CCTTCTTCCCTTCTCTC 1757
1017 TTGAAGATATTTACTTGGCCCTTTAAAGTTGGGAATCTTCACTCTTCTATACCTATTATC 958
1758 TATTCCTTTGGTTTGTGATAGTGTCTCTGGCTTCCCTGGAGTTTATGCTCGGATTAAT 1817
957 CTAGTTTCACTTCTCACTTGTCTCTGAATTTCTTGGATGATTTGGGTTAGATCTTT 898
1818 TTAGACTTAAATTTCTTTTACCAAGGTATCCATTTCTTCTATCTTCTTCTCACTGCT 1877
897 TTGCTTTTGGCATTTCTTTCACCTGTTGTGT-CAGTGTCTTATGATATCTTCTTCACT 839
1878 GAGATCTCTCTCTCTCTCTTGTATCTGTGTCAGTGGCTTGTCTGAGGTTCTGAGGTTCTG-T 1936
838 GAGATCTCTCTCTCTCTCTTGTATCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 779
1937 TGGGTTCTTAAATTTTCTTCAATTTCCAGATTTCTTCACTTGGGTTTGGTGTATTAATTTCT 1996
778 TCTTCTCTAGTTTCTTCAACCCAGGTTGTCTCTCTTGTGTTGTTGTTGTTGTTGTTGTT 719
1997 ATTCCACTTTCAGTCTCTGAAATGTTTACTCAATTTCTTCTTCTTCTTCTTCTTCTTCT 2053
718 ATTCCATTTTGTAGATCTCGATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 659
2054 TCATAGTTTCTTAAATGCAATTTTCAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2095
658 TCTGTATTTCTTAAAGGATTTGTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 617

BZ097284 835 bp DNA linear GSS 10-OCT-2002
CH230-236115-TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-236115, genomic survey sequence.

ACCESSION
VERSION
BZ097284.1 GI:23738168

KEYWORDS
GSS.

SOURCE
Rattus norvegicus (Norway rat)

ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
1 (bases 1 to 835)
Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,
Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
Riggs, F., de Jong, P. and Fraser, C.M.
Rat BAC End sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other GSSs: CH230-236115.TU
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research

TITLE
JOURNAL
COMMENT

1232-TGGAATTTATTTATTTCAATTTTCTGTAATGCGGTGAACATCTTTAGATTGAAGTTT 1291
1495 TTGAAGATTACTTCTTCTGTTTCAAGGCTGTAGTTTGTCTTCTTATAGCTGTTTCTC 1436
1292 TCTCTAGCTTTTGTAGTCTGATTTGAAGTAGATATCTTCACTGATTTTATC 1351
1435 ATCTATTATCTTTGTAGGCTCGATTTGTAAGAGATATGTTGTAATTTGGATTGTTC 1376
1352 TTGAATGTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1411
1375 ATGAATACITTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1316

Query Match 15.2% Score 347.299; DB 1; Length 2559;
Best Local Similarity 66.28; Pred No. 6.8;
Matches 584; Conservative 0; Mismatches 277; Indels 21; Gaps 6;

misc_feature 1..2559
/notes="similar to PORFI [Mus musculus domesticus]"
(SPTRIQ91V68, evidence: FASTV, 70.8%ID, 100%length,
match=1094)"

FEATURES
source
1..2559
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430016D20"
/db_xref="MGI:2427317"
/db_xref="taxon:10090"
/clone="E430016D20"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"

Location/Qualifiers
1..2559
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430016D20"
/db_xref="MGI:2427317"
/db_xref="taxon:10090"
/clone="E430016D20"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

location/Qualifiers
1..2559
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430016D20"
/db_xref="MGI:2427317"
/db_xref="taxon:10090"
/clone="E430016D20"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"

misc_feature 1..2559
/notes="similar to PORFI [Mus musculus domesticus]"
(SPTRIQ91V68, evidence: FASTV, 70.8%ID, 100%length,
match=1094)"

Query Match 15.2% Score 347.299; DB 1; Length 2559;
Best Local Similarity 66.28; Pred No. 6.8;
Matches 584; Conservative 0; Mismatches 277; Indels 21; Gaps 6;

1232-TGGAATTTATTTATTTCAATTTTCTGTAATGCGGTGAACATCTTTAGATTGAAGTTT 1291
1495 TTGAAGATTACTTCTTCTGTTTCAAGGCTGTAGTTTGTCTTCTTATAGCTGTTTCTC 1436
1292 TCTCTAGCTTTTGTAGTCTGATTTGAAGTAGATATCTTCACTGATTTTATC 1351
1435 ATCTATTATCTTTGTAGGCTCGATTTGTAAGAGATATGTTGTAATTTGGATTGTTC 1376
1352 TTGAATGTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1411
1375 ATGAATACITTTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1316

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1. (bases 1 to 838)
Zhao,S., Shetty,J., Shatsman,S., Teegayee,G., Geer,K.,
Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.
Rat BAC End sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other GSSs: CH230-232F19.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
7912 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(<http://www.chori.org/bacpac/oreringinformation.htm>). BAC end
page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html
Plate: 232 row: F column: 19
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. .838
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SENHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-232F19"
/sex="Female"
/cell_type="Brain"
/clone_lib="CHORI-230 Segment 1"
/note=Vector: pPARBAC2.1; Site_1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by
Pieter de Jong"

FEATURES
source

Query Match 15.1%; Score 345.099; DB 1; Length 838;
Best Local Similarity 67.1%; Pred. No. 20;
Matches 557; Conservative 0; Mismatches 254; Indels 19; Gaps 5;

Qy 1213 CCTCTTTGATTTTGGCGCTGGAAATATTATTATTCATATTTTCTTGAATGTGGGTAA 1272
Db 836 CATCTCTGTGTTGCGCCCAAGACGATTAGTCTGCTTTTCTCTAGGGTATAGCTGCC 777
Qy 1273 ATCTTTAGATTGAAGTTTCTTCTCTAGCCTCTT--TAGTCTGCATTTGAAGATAGATA 1330
Db 776 CTCCTTATGTGGGCTTTACCCTTTATATCCTTTGTAGTGTGGATTGTGAGAAAGATA 717
Qy 1331 TTCTTTACATCTGATTTTATCTTAGAATGCTCTTCTTCTCCAACTATTGTGACAGAAAG 1390
Db 716 TTGCTAAATTTGGTTTGTGATGGAATATCTTGTTTCCATCTATGTTAATTGAGAG 657
Qy 1391 TTTTCTTAAGTGCAGTAGTCTGGCGCTGCACATCTCTAGTCTCTGGAGTCTGTAGCACATC 1450
Db 656 TTTTGCAGGATACAGTAACTCGGCTGGCATTTGTTCTCTTAGGCTCTGATGACATC 597
Qy 1451 TGTGAGGCGCTTCTTACATTTTGAGTTTCTATTGGAAAGTCAGGTGTAACTCTAATAC 1510
Db 596 TGTCCAGATCTTCTGGGCTTCATAGTTTCT--GSCGAAAGTCTGTGTGATCTGTATAG 538
Qy 1511 ATCTGCCCTTATATGTTAAATGGTCTTTTCTCCCTTGCATCTTTTAATATCTTCTTTG 1570
Db 537 GTCGCTCTTATATGTTACTT-GACCTTTTCCCTTACTGCTTTTAATATCTTCTTTTA 479
Qy 1571 TTCTATATCTTTTAGTGATTTCGATTATATGCACTGTGGGAGTTCTCTTTCGGTCCAAAT 1630
Db 478 TTTCTGTGTGTGGTGTGTTTGACATTAATGACGGTAGGTGTTCTTTCTGCTCACT 419
Qy 1631 CTATTTGGTGTTTGTATGCTCTCTGTACCTGTATGAGCACTCTCTTCTCAAGTTAGGA 1690
Db 418 CTATTTGAGTCTCTAGGCTTCTGTATGTCTATGGGTATCTCTTTTGTAGTTAGGG 359

Qy	1691	AAATTTTCCTTTTGGTTTCTTGAAATAATTTTCCCTGCTTTTGACCTG-----	1744
Db	358	AAGTTTCTTCTTAAGATTTTGTGAAGATATTTACTGGTCTTTTGACGTGGAGTCTTCA	299
Qy	1741	-----CCTTCTCCCTTCCCTATATCCTTTGGTTTTTGCATAGTGTCTCGGCTTCCCTG	1796
Db	298	CTCTCTTCTATACCTATATACCTTAGTTTGAATCTCTCAATGAGTCCTGGAATTCCTGT	239
Qy	1797	ATGTTTATGCTGGATATTTTAGACCTTAACATTTTCTTTTGACCAAGGTATCAATTTCT	1856
Db	238	ATGTTTGTGGACCAAGTACCTTTTCCGCTTACAAATATCTTTGACGTGAGTCAATGATT	179
Qy	1857	TCTATCTTGCTTCACTGCTCGAGATCTCTCTTCTATCTCTTGTATCTCTGTCAGTGAGG	1916
Db	178	TCTATGAATCTTCTGCTCTCTGAGATCTCTCTCCATCTCTGTATCTGTGGTGAAG	119
Qy	1917	CTTGCTCTGAGGTCTCTGTGGGTCTTAATTTTTTC-ATTTCCAGATTTCTCTCAGTT	1975
Db	118	CTTGATCTACAGCTCCTTCTCTTTCTTTTGTGATTTTCTATGTCGAGGTTGTTCCATG	59
Qy	1976	TGGTTTTTGTATTAACTTATTTCCATTTCCAGTTCCAGTCTCGAAATGTTTT	2025
Db	58	TGTTCTTCTTGATCTGCTCTATTTCCATTTTAACTTCTTCACTGTTTT	9
RESULT	42		
LOCUS	BH258120/c		
DEFINITION	CH230-112118.TU CHORI-230 Segment 1 Rattus norvegicus genomic clone		
ACCESSION	BH258120		
VERSION	BH258120		
KEYWORDS	GENS.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
AUTHORS	1 (bases 1 to 849) Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K., Shwartsbeyn,A. Georgeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de Jong,P. and Fraser,C.M.		
TITLE	Rat BAC End Sequences from Library CHORI-230 EcorI segment		
COMMENT	Unpublished (1999) Other GSSs: CH230-112118.TV Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: szhao@tigr.org Clones are derived from the rat BAC library CHORI-230 (http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/or ering_information.htm). BAC end plate: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html Page: 112 row: 1 column: 18 Seq primer: SP6 Class: BAC ends		
FEATURES	Location/Qualifiers		
source	1..849 /organism="Rattus norvegicus" /mol_type="genomic DNA" /strain="BN/SENHsd/MCW" /db_xref="taxon:10116" /clone="CH230-112118" /sex="Female" /cell_type="Brain" /clone_lib="CHORI-230 Segment 1" /note=Vector: pTARBAC2.1; Site_1: EcorI; Site_2: EcorI; CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by		

Pieter de Jong

Query Match 15.1%; Score 345.099; DB 1; Length 849;
 Best Local Similarity 66.0%; Pred. No. 19;
 Matches 555; Conservative 0; Mismatches 269; Indels 17; Gaps 4;

QY 1201 TGTGTTCTCCCTCTTTGATTTTGGCTCGGAATATTATTATTCATATTTCTTG 1260
 DB 841 TTTTGGGCTTTCATCTCTTTGTTGGTCCCAAGATGATTAGTTCTGTTCTCTAG 782
 QY 1261 AATGTGGTAAACATCTTTAGATTGAAGTT-TTCTCTAGCCCTCTTTAGGCTCGCATTT 1319
 DB 781 GGTATAGCTGCCCTCCCTATGTTGGCTTTACCAATTTATATCTTTGTAGTGTGATTT 722
 QY 1320 GAAGATAGATATTCTTTACATCTGAATTTATCTTAGAATGTTCTTTCTCCCACTATT 1379
 DB 721 GTAGAAAGATATTGTGTAATTTGGTTTGTATGGAATATCTTGGTTTCTCCATCAATG 662
 QY 1380 GTGACAGAAAGTTTCTAAGTCAGTAGTCTGGCTCGACATCTAGTCTCTTGGAGTC 1439
 DB 661 TTAATTCAGAGTTTGTCTGATACAGTAACCTGGCTGGCAATTTGTCTCTTAGGGTC 602
 QY 1440 TGTAGCACATCTGTGAGGCTCTTTACATTTTGTAGTTTCTATTTGAAAGTCAGGTGT 1499
 DB 601 TGTATGATATCAGTCCAGGATCTTCTGGCTTCATAGTTTCT-GGCGAGAGTCTGGTCT 543
 QY 1500 AATCTAATACATCTGCTTTATATGTTAATGGTCTTTTCCCTTGCATCTTTTAATA 1559
 DB 542 GATCTGATAGGCTGCTTTATATGTTACTTGACCTGTTTCCCTTACTGCTTTTAATA 483
 QY 1560 TTCTTTCTTTGTTCTATPACTTTAGTGAATTAATATGACATCTGGGGAGTTCTTTT 1619
 DB 482 TTCTTTCTTTTGTGCGTTTGTGTTGACAAATTAATGACGGAGGTTTCTTTT 423
 QY 1620 TCGGGTCAATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1679
 DB 422 TCTGGTCCAACTATTTGGAGTTCTGAGGCTTCTGTATGTTCTATTTGTTATCTCTTTT 363
 QY 1680 CAAGGTAGGAATTTTCTTTTGTGTTTCTGAAATATTTCCCTGCTTTGACCT 1739
 DB 362 TTAGGTTAGGAAGTTTCTTTCTATGATTTTGTGAAGATATTTACTGGTCTTTGAGCT 303
 QY 1740 G-----CCTTCTCCCTTCTCTATTTCTTCTTTTGGTTTGTGCATAGTGTCTC 1785
 DB 302 GGGAGTCTTCTCTCTCTATACCTATTATCTTAGTTTGTATCTTCTCATGTGCTCT 243
 QY 1786 TGGCTTCTGGATGTTTATGCTCGGATATTATTAGCATTAACATTTCTTTGACCAAGG 1845
 DB 242 GGATTTCTGTATGTTTGGACAGTAGCTTTTCTGCTTTACATTTATCTTTGACAGTTG 183
 QY 1846 TATCCATTTCTTCTATCTCTCTCACTGCTGAGATCTCTCTTCTATCTCTGTATTTC 1905
 DB 182 AGTAATATTTCTATGGAATCTTCTGCTCTGAGATTTCTCTTCCATCTCTGTATTC 123
 QY 1906 TGTCAGTAGGCTTGTCTGAGGTTCTGTGTTGGGTTCTTAATTTTTC-ATTCCAGAT 1964
 DB 122 TGTGTGGAAGCTGTATCTACAGCTCCCTGTCTCTTTCTTTTGGTTTCTATATCCAGG 63
 QY 1965 TTCTTTCAGTTTGGGTTTGTATTATTAATTTCAATTTCCACTTTCCAGTCTCGAATGTTT 2024
 DB 62 TTGTTTCCATGTTTCTTTCTTGAATGCTTCTATTTCCAAATTTAATCCCTCACTGTT 3
 QY 2025 T 2025
 DB 2 T 2

RESULT 43

BH327408/c

LOCUS

DEFINITION

CH230-174G8.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone

CH230-174G8, genomic survey sequence.

ACCESSION

BH327408

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BH327408.1 GI:17259122

GSS.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 821)

Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K.,

Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,

Riggs, F. de Jong, P. and Fraser, C.M.

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

Unpublished (1999)

Other GSS: CH230-174G8.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(http://www.chori.org/bacpac/rat230.htm). For BAC library

availability, please contact Pieter de Jong (pjejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or_ering_information.htm). BAC end

page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html

Plate: 174 row: G column: 8

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..821

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SENHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-174G8"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 1"

/note=vector: pTARHAC2.1; Site 1: EcoRI; Site 2: EcoRI;

CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by

Pieter de Jong"

Query Match 15.1%; Score 344.999; DB 1; Length 821;

Best Local Similarity 67.0%; Pred. No. 20;

Matches 544; Conservative 0; Mismatches 250; Indels 18; Gaps 4;

QY 1224 TTTTGGCTGGAAATTTATTATTATTCATATTTTCTTGAATGTGGGTAACTTTTGAATT 1283

DB 810 TTTTGGCTGGAAATTTATTATTATTCATATTTTCTTGAATGTGGGTAACTTTTGAATT 751

QY 1284 GAAGTTTTC--TCCTAGCCTTCTTTAGGCTGCAATTTGAGATAGATATTTTACATC 1341

DB 750 GGTCTTTTACCATTATTCATCTTTGTAGTCTGGATTTGTAGAAAGATATTGTGAAAT 691

QY 1342 TGATTTTATCTTGAATGTCTTTCTTTCTCCAACTATTGTGACAGAAAGTTTCTTAAAGT 1401

DB 690 TGGTTTGTCTATGGAATATATTGGTTTCTCCATCAATGTTAACTGAGAGTTTGTCTGAT 631

QY 1402 GCAGTAGCTGGCCCTGACATCTGTAGTCTCTTGGAGTCTGTAGACATCTGTGACGGCC 1461

DB 630 ACAGTAACTCTGGATGGCATTTGTGTTCTCTTAGGCTCTGTATAAATCATCATGTCAGGATC 571

QY 1462 TTCTTACATTTTGAATTTCTATTGGAAAGTCAAGGTGAATTTCTTAATATCATCTGCTTTA 1521

DB 570 TTTTGGCTTCATAGTTTCT--GGCGAGAAGTCTGGTGTGATTTCTTATAGTCTCCCTTTA 512

QY 1522 TATGTTAATTTGGTCTTTTTCCTTTCCTTGCATCTTTTAAATATTTCTTTTGTCTATCTTT 1581

DB 511 TATGTTACTT-GACCTTTTTCCTTCTACTGCTTTTAGTATTTCTTTTATTTTGTGCTT 453

QY 1582 TAGTGATTTGATTTATGACATCTGCGGAGTTTCTTTTCCGGTCCCAATCTATTGTGTG 1641

DEFINITION CH230-163N10.TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-163N10, genomic survey sequence.

ACCESSION BH357910

VERSION BH357910.1 GI:17288644

KEYWORDS GSS.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 762)

Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K.,
Shvartsbeyn,A., Gebregorgis,E., Overton,L., Russell,D., Chen,D.,
Riggs,F., de Jong,P. and Fraser,C.M.

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

Unpublished (1999)

Other_GSSs: CH230-163N10.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.choi.org).

Clones may be purchased from BACPAC Resources

(<http://www.chori.org/bacpac/oreringinformation.htm>). BAC end

page: http://www.tigr.org/tldb/bac_ends/rat/bac_end_intro.html

Plate: 163 row: N column: 10

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..762

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SsNHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-163N10"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 1"

/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;

CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by

Pieter de Jong"

Query Match 15.1%; Score 343.899; DB 1; Length 762;

Best Local Similarity 70.2%; Pred. No. 22;

Matches 526; Conservative 0; Mismatches 206; Indels 17; Gaps 5;

1301 CTTCTTTAGCTGCAATTCAGATAGATATCTTTACATCTGATTTATCTTAGAATGT 1360

753 CTTTCTAGCGCTGGATTTGTAGAAAGATATTGTGTAATTTGGTTTGTACGGAATAT 694

1361 CTTTCTTTCTCCCACTATTGTGACAGAAAGTTTCTTAAGTGCAGTAGTCTGCGCTGACA 1420

693 CTTAGTTTCTCCATCAATGTTAATTGGGAGTTTTCAGGATACAGTAACCTGGGCTGCA 634

1421 TCTGTAGTCTTTGGAGTCTGTAGACATCTGTGACAGGCGCTTCTTACATTTTGAGTTTC 1480

633 TTTGTGTCTCTTAAGTCTGTATGACATCTGTCCAGGAACATCTGGCTTTCATAGTCTC 574

1481 TATTGGAAGTCAGGTAAATCTTAATACATCTGCTTTATATGTTAATTTGGTCTTTT 1540

573 T-GGAGAGAGTCTGGTGTGATCTCTGATAGGCTGCTTTATATGTTACTT-GACCTTTT 516

1541 TCCTTTGCATCTTTTAAATATCTTTCTTTGTTCTATATCTTTTAGTGATTTGATTTATG 1600

515 TCCTTTACTGCTTTTAAATATCTTTCTTTTATTTGAGCATTGATGTTTGGACATTATG 456

1601 CACTGTGGGGAGTTCTTTTCCGGTCCCAATCTATTGGTGTGTTTGTATGCTTCTTGTTACC 1660

Db 455 TGATGGGGGAGTTTCTTTTCTGGTCCCAATCTATTGGAGTCTGTAGGCTTCTTGTATG 396

Qy 1661 TTGATAGGCATCTCTTTCTCAAGGTAGGAATTTTCTTTTGGTTTCTTGAATAAATA 1720

Db 395 TTTATGGGCATCTCTTTCTTTAGGTAGGGAAGTTTCTTCTATGATTTTGTGAGATA 336

Qy 1721 TTTTCCCTGCTTTTGACCTGC-CTTCTTCCCTTCTCTATTC------TTT 1766

Db 335 TTTACTGCTCTTTTGAGCTGGCGCTTCTCACTCTCTCTATACCTATATCCITTAGGTTT 276

Qy 1767 GGTTTTTCATAGTGTCTCTGGCTTCTCGGATGTTTATGCTCGATATTTTATAGACTTA 1826

Db 275 GATCTTCTCATGAGTCTCTGGATTTTCTGTATGTTTGGACCACTAGCTTTTCTGTTT 216

Qy 1827 ACATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTTGTCTTCTCACGCTGAGATTCTC 1886

Db 215 ACATTAICTTTGACAGTTGTGTCGATGATTTCTAAGGAATCTTCTCTCCGAGATTCTC 156

Qy 1887 TCTTCTATCTCTGTATTTCTGTCAAGTGGCTTGTCTCTGAGTTTC- TGTGGGTTCTT 1945

Db 155 TCTTCTATCTCTGTATTTCTGTGTAATGCTTGTATCTATGGCTCTTGTCTCTTCTT 96

Qy 1946 AATTTTTCATTTCCAGATTTCTTCAAGTTGGGTTTGTGTTTATTAATTTCTATTTCCACT 2005

Db 95 TGGTTTCTATATCCAGGTTTGTTCCTTTGTGCTTTCTTATGCTTCTATTTCCATT 36

Qy 2006 TTCAGGTCCTGAAATGTTTACTCATTTT 2034

Db 35 TTTAATCTCTCACTGTTTGGCTTGTGTT 7

Search completed: August 25, 2004, 09:33:44

Job time : 181 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 08:45:30 ; Search time 14588 Seconds
(without alignments)
3.946 Million cell updates/sec

Title: US-10-664-775-2

Perfect score: 3572

Sequence: 1 gtcaggagggcgcagtgta.....gcacacacgcagaagctt 3572

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 44 seqs, 8057215 residues

Total number of hits satisfying chosen parameters: 88

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : rge2.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	2197.37	61.5	160990	1	AC141413
C 2	2158.67	60.4	233345	1	AC141481
C 3	2153.4	60.3	198916	1	AC145008
C 4	2147.7	60.1	179015	1	AC135839
C 5	2147.7	60.1	241923	1	AC137656
C 6	2131	59.7	173068	1	AC091660
C 7	2125.5	59.5	171153	1	AC135215
C 8	2123.49	59.4	145088	1	AC142564
C 9	2123.49	59.4	192431	1	AC141853
C 10	2108.1	59.0	173207	1	AC108887
C 11	2107.7	59.0	213549	1	AC140963
C 12	2106.5	59.0	196994	1	AC144887
C 13	2105.2	58.9	192199	1	AC139312
C 14	2103.2	58.9	145088	1	AC142564
C 15	2100	58.8	189278	1	AC142562
C 16	2095.4	58.7	224721	1	AC139632
C 17	2081.09	58.3	193363	1	AC091252
C 18	2071.6	58.0	139617	1	AC146889
C 19	2071.2	58.0	172022	1	AC145440
C 20	2066.1	57.8	294688	1	AC144498
C 21	2065.3	57.8	153264	1	AC135216
C 22	2065	57.8	223761	1	AC147195
C 23	2062.1	57.7	180432	1	AC134959
C 24	2052.6	57.5	185590	1	AC146537
C 25	2046.4	57.3	181921	1	AC091399
C 26	2032.6	56.6	175949	1	AC107065
C 27	2032.6	56.5	177535	1	AC109914
C 28	2018.9	55.5	207875	1	AC147592
C 29	1982.6	54.5	192540	1	AC137651
C 30	1945.1	54.5	192540	1	AC137651
C 31	1931.7	54.1	163604	1	AC092727
C 32	1930.2	54.0	168497	1	AC092410
C 33	1919.1	53.7	173952	1	AC130787

ALIGNMENTS

```

RESULT 1
AC141413/c
LOCUS      AC141413      160990 bp      DNA      linear      HTG 07-MAY-2003
DEFINITION Bos taurus clone RP42-99E14, WORKING DRAFT SEQUENCE, 8 ordered
            pieces.
ACCESSION  AC141413
VERSION    AC141413.2 GI:30409902
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE     Bos taurus (ccw)
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
REFERENCE  1 (bases 1 to 160990)
            Antonellis A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
            Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K.,
            Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,
            Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,
            Hurler, B., Idol, J.R., Karlins, E., Kwong, P., Lalic, P., Lee-Lin, S.-Q.,
            Legaspi, R., Maduro, Q., J., Maduro, V.B., Margulies, E.H., Masello, C.,
            Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E.,
            Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Shah, K.,
            Sison, C., Stantripp, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,
            Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
            NISC Comparative Sequencing Initiative
            Unpublished
            Green, E.D.
REFERENCE  2 (bases 1 to 160990)
            Direct Submission
            Submitted (14-MAR-2003) NIH Intramural Sequencing Center, 8717
            Grovmont Circle, Gaithersburg, MD 20877, USA
REFERENCE  3 (bases 1 to 160990)
            Direct Submission
            Submitted (07-MAY-2003) NIH Intramural Sequencing Center, 8717
            Grovmont Circle, Gaithersburg, MD 20877, USA
            On May 7, 2003 this sequence version replaced gi:28951185.
            ----- Genome Center
            Center: NIH Intramural Sequencing Center
            Center code: NISC
            Web site: http://www.nisc.nih.gov
            Contact: nisc.zoo@nih.gov
            ----- Project Information
            Center project name: edv
            Center Clone name: 099E14

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence

contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 159861 bases at least Q40
 Consensus quality: 160028 bases at least Q30
 Consensus quality: 160216 bases at least Q20
 Insert size: 162000; agarose-fp
 Insert size: 160290; sum-of-contigs
 Quality coverage: 11.34x in Q20 bases; agarose-fp
 Quality coverage: 11.46x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 50852: contig of 50852 bp in length
 50853 50952: gap of unknown length
 50953 87597: contig of 36645 bp in length
 87598 87697: gap of unknown length
 87698 90744: contig of 3047 bp in length
 90745 90844: gap of unknown length
 90845 97904: contig of 7060 bp in length
 97905 98004: gap of unknown length
 98005 106019: contig of 8015 bp in length
 106020 106119: gap of unknown length
 106120 130293: contig of 24174 bp in length
 130294 130393: gap of unknown length
 130394 156781: contig of 26388 bp in length
 156782 156882: gap of unknown length
 156882 160990: contig of 4109 bp in length.

FEATURES

source

1. .160990
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-59814"
 /clone_lib="RP42"

misc_feature

1. .50852
 /note="assembly_fragment
 clone end:SP6
 vector_side:left"

misc_feature

50953..87597
 /note="assembly_fragment"
 87698..90744

misc_feature

/note="assembly_fragment"
 90845..97904

misc_feature

/note="assembly_fragment"
 98005..106019

misc_feature

/note="assembly_fragment"
 106120..130293

misc_feature

/note="assembly_fragment"
 130394..156781

misc_feature

/note="assembly_fragment"
 156882..160990
 /note="assembly_fragment
 missing T7 clone end on 3' end of insert"

Query Match

Best Local Similarity 61.5%; Score 2197.3; DB 1; Length 160990;
 Matches 2635; Conservative 77.5%; Pred. No. 0;
 Mismatches 162; Indels 601; Gaps 4;

QY 1 GTCAGGAGGGCGGCGAGTGGAGAGGTACCTACCTCGTCCAGGTAAAGGACGATGCG 60

DB 115101 GTCAGGAGGGTGGCGGTAGGAGA-----TACCCTCGTCCAGGTAAAGTACGACGC 115047

QY	61	TGCGCTTGTCTGGAGCAGCGCTAAAGAGATACCCACGCCCAAGGTAAAGAAACCCCAAG	120
DB	115046	TGCCCTTGTCTGGAGCAGCGCTAAAGAGATACCCCAAGGTAAAGAAACCCCAAG	114987
QY	121	TAAGATGGTAGGTGTGTGAGAGGCGCATCAGAGGCGAGACATCTCTGAAACCATACAGCA	180
DB	114986	TAAGACGGTAGGTGTGCAAGAGGCGCATCAGAGGCGAGACATCTCTGAAACCATACAGCA	114927
QY	181	GAATACTAGTCAATCTAATCACTAGGACACACAGCCTTGTCTAACTCAATGAACCTAAG	240
DB	114926	GAATACTAGTCAATCTAATCACTAGGACACACAGCCTTGTCTAACTCAATGAACCTAAG	114867
QY	241	CCATGCCCGTGGGGCAACCCAAAGATGGGCGAGTCAATGCTGGAGAGATCTGACGAATGTG	300
DB	114866	CCATGCCCTGTGGGGCAACCCATGATGGGGGGTCAATGCTGGAGAGATCTGACGAATGTG	114807
QY	301	GTCCACTGGAGAGGGAATGCAAACTTCACTAGTATTTCTTGCCTTGAACCCCATGAAC	360
DB	114806	GTCCACTGGAGAGGGAATGCGAAACCACTTCACTAGTATTTCTTGCCTTGAACCCCATGAAC	114747
QY	361	AGTATGAAGGCGCAAAATGATAGGATCTGAAAGAGGAACCTCCACAGTCTAGTAGTGCC	420
DB	114746	AGTATGAAGGCGCAAAATGATAGGATCTGAAAGAGGAACCTCCACAGTCTAGTAGTGCC	114687
QY	421	CCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCA	480
DB	114686	CAATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCA	114627
QY	481	RAGCAAAAGAAATACCCAGCTGTGGATGTGATGCTGGTGTATATAGCAAGGTCCGATGCTGT	540
DB	114626	RAGCAAAAGAAATACCCAGCTGTGGATGTGATGCTGGTGTATATAGCAAGGTCCGATGCTGT	114567
QY	541	AAAGAGCAATATTGTCATAGGAACCTGGAATCTCAGTCCATCAATCAAGGCAAAATGGAA	600
DB	114566	AAAGTGCATATTGTCATAGGAACCTGGAATCTCAGTCCATCAATCAAGGCAAAATGGAA	114507
QY	601	GTGCTCAAAAGAGATGGCAAGATGAATGTCAACATTTAGGAATCAGCGAACTAAAA	660
DB	114506	GTGCTCAAAAGAGATGGCAAGATGAATGTCAACATTTAGGAATCAGCGAACTAAAA	114447
QY	661	TGGACTGGAATGGGTGAATTTAACTCAGATGACCATATATCTACTCTCGGCGACGAA	720
DB	114446	TGGACTGGAATGGGTGAATTTAACTCAGATGACCATATATCTACTCTCGGCGACGAA	114387
QY	721	TCCCTCAGAGAAATGGAGTAGCCATCATGTCACCAAAAGAGTCCGAAATGCACTACTT	780
DB	114386	TCCCGCAGAGAAATGGAGTAGCCATCATGTCACCAAAAGAGTTCGAAATGTAGCACTT	114327
QY	781	GGATGCAATCTCAAAAACGACAGATGATCTCTGTTTCTCCAGGCAACCAATTCAT	840
DB	114326	GGATGCAATCTCAAAAACGACAGATGATCTCTGTTTCTCCAGGCAACCAATTCAT	114267
QY	841	ATCAGATTAATCCAAAGTCTATGCCCCAACCAAGTAACTGCTGAAGAGGTGAAGTTGAACGG	900
DB	114266	ATCAGATTAATCCAAAGTCTATGCCCCAACCAAGTAACTGCTGAAGAGGTGAAGTTGAACGG	114207
QY	901	TCCCTATGAGACCTTACAGACCTTTTAGAATAACACCCCAAAAAGATGCTCTTCTCAT	960
DB	114206	TTCCTATGAGACCTTACAGACCTTTTAGAATAACACCCCAAAAAGATGCTCTTCTCAT	114147
QY	961	ATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAACACCTGGAGTAAACAGGCAAAATTT	1020
DB	114146	ATAGGGGACTGGAATGCAAAAGTAGGAAGTAAAGAACCTGGAGTAAACAGGCAAAATTT	114087
QY	1021	GGCTTGAATACCGAATGAAGAGGGGCAAGACTAATAGAGTTTTGCCAAGAAATGCA	1080
DB	114086	GGCTTGAATACCGAATGAAGTAGGGCAAGACTAATAGAGTTTTGCCAAGAAATGCA	114027
QY	1081	CTGTGCTAGCAACACCCCTCTTCCAAACACACAAAGAGAGACTCTACACATGACATCA	1140
DB	114026	CTGTGCTAGCAACACCCCTCTTCCAAACACACAAAGAGAGACTCTATACATGACATCA	113967
QY	1141	CCAGATGGTCAACACCGAAATTCAGATTTGATTTATTTCTTTCAGCCCAAGATGGGAAGC	1200

113966 CCAGATGGTCAACACCGAAATCAGATTGATTATATCTTCGACGCAAGATGGAGAGC 113907
1201 TCTATACAGTCAAGCAAAACAGACACAGAGCTTACTGTGGCTCAGATCATGAATCCTT 1260
113906 TCTATACAGTCAAGCAAAACAGACACAGAGCTGACTGTGGCTCAGATCATGAATCCTT 113847
1261 ATTGCCAAATTCAGACTTAAATTTGAAGAAAGTAGGGAAGAAACCACTAGATCATCAGGTAA 1320
113846 ATAGCCAAATTCAGACTTAAATTTGAAGAAAGTAGGGAAGAAACCACTAGATCATCAGGTAT 113787
1321 GACCTAAATCCAAATCCCTTATCATATATACAGTGAAGTAGAGAAATAGATTTAAAGGCCCTA 1380
113786 GACCTAAATAAATTTCCCTTATCATATATACAGTGAAGTAGAGAAATAGATTTAAAGGCCCTA 113727
1381 GATCTGATAGACAGAGTACTTAATGAACATATGACAGAGAGTTTCATGATGTACAGAG 1440
113726 GATCTGATAGATAGAGTACTTGAATGAACTATGGAATGAGTTTCGTCGATGTATATGAG 113667
1441 ACAGGATCGACACCATCCCAATGGAAGAAAGAAATGCAAAAGCAAAATGGCTGTCTGG 1500
113666 ACAGGATCAAGACCATCCCAATGGAAGAAAGAAATGCAAAAGCAAAATGGCTGTCTGG 113607
1501 GAGGCCCTTACAAATAGCTCTGTAAGAAAGAGAGAGTGAAGCAAAAGGAAAGAAAGAA 1560
113606 GAGGCCCTTACAAATAGCTCTGTAAGAAAGAGAGAGTGAAGCAAAAGGAAAGAAAGAA 113547
1561 TAAAGCATCTGAATCAGAGTTCCAAA 1588
113546 TATAAGCATCTGAATCAGATTTCCTCAAGAAATAGCAAGAGAGATGAAGAAAGCTTCCTC 113487
1589 1588
113486 AGCGATCAATGCAAGAAATAGAGGAAACACACAGNATGGGAAAGATAGAGATCTCTTC 113427
1589 1588
113426 AAGAAAATAGAGATACCAAGGAACATTTTCATCAAGATGGCTCGATAAAGACAGA 113367
1589 1588
113366 AATGGTAGGACCTAACAGAGCAGAGATATGAGAGAGATGGCAAGAAATACACAGAA 113307
1589 1588
113306 GAATGTACAAAAGATCTTACAGAGCCAGATAATCACGATGGTGTGATCACTGACCTA 113247
1589 1588
113246 GAGCCAGACATCTTGGAAATGTGAAGTCAAGTGGCCCTTAGAAAGCATCTACAAACAAA 113187
1589 1588
113186 GCTAGTGGAGTGAATTCAGGTGAGCTATTTCAATCCTGAAAGATGATGTGTG 113127
1589 1588
113126 AAAGTGTGCATCAATATGCCAGCAAAATCTGGAAAACCTCAGCAGTGGCCACAGGACTGG 113067
1589 1588
113066 AAAAGGTCAATTTCAATTCOAATCCCAAGAAAGGCAATCCCAAGAAATGCTCAAACTAC 113007
1589 1588
113006 GGCACAGTTGCATCTCACACGCTAGTAAAGTAATGCTCAAAATTTCTCCAAGCCAGG 112947
1589 1588
112946 CTTCAGCAAGATTTGAAGTGTGAATCTTCTGATGTTCAAGCTGTTTTAGAAAAGGAGA 112887
1628 GGAACGAGACCAAAATGCAACATCTCTGTATCATGGAAGAAAGCAAGAGATTCAG 1687

112886 GGAACGAGAGATCAAAATTCGCAACATTTGCTGGATCATAGAAAAAGCAAGAGAGTTCCAG 112827
1688 AAAAAATCTATTTCTGCTTTTATGACTATGCAAAAGCCTTTTACTGTGGGGTCACAAT 1747
112826 AAAAAATCTATTTCTGCTTTTATGACTATGCAAAAGCCTTTTATGTTGTGGATCACAAT 112767
1748 AACTGTGGAAAAATTTCTGAAAGGATGGGAATACCAAGACCACCTGATCTCCCTCTTGAAA 1807
112766 AACTGTGGAAAAATTTCTGAAAGGATGGGAATACCAAGACCACCTGATCTCCCTCTTGAGA 112707
1808 AATTGTATCGAGTCAAGGAAAGCAACGTTAGAACTGGACATGGAACCAAGAGATGGTTTC 1867
112706 AATCTATATGAGTCAAGGAAAGCAACGTTAGAACTGGACATGGAACCAAGAGATGGTTTC 112647
1868 CAAGTCAAGAAAGGAGTATGTCAGGCTGTATATTTGTCACCCGCTGTTTAACTTCTAT 1927
112646 CAATAGAAAAGGAGTACGTCAGGCTGTATATTTGTCACCCCTGTTTAACTTATAT 112587
1928 GCAGAG-ACATCATAGAAAACGCTGGCTGGAAAGAACACAAGCTGGAATCAAGATTGCC 1986
112586 GCAGAGTACATCATGAGAAAACCTGGACTGGAAGAAAACAAGTTGGAATCAAGATTGCC 112527
1987 GGGAGAAATAGCAATAACCTCAGATATGCAGATGATACCAACCTTATGGCAGAAAAGTGA 2046
112526 GGGAGAAATATCAATAACCTCAGATATGCAGATGATACCAACCTTATGGCAGAAAAGTGA 112467
2047 GAGAACTTAAAAAGCCTCTTTCATCAAGGTGAAAGAGGAGAGTGAAGAGTTGGCTTAAAG 2106
112466 GAGAACTTAAAAAGCCTCTTTCATGAAAGTGAAGTGGAGAGTGAAGAGTTGGCTTAAAG 112407
2107 CTCAACTTCAAGAAAAGAGATCATGGCATCTGGTCCCATCACTTTCATGGGAATAGAT 2166
112406 CTCAACTTCAAGAAAAGAGATCATGGCATCTGGTCCCATCACTTTCATGGGAATAGAT 112347
2167 GGGAAAAACAGTGGAAAAACAGTGTACAGCTTTTATTTTGGGGGGCTCCAAAATCACTGCAG 2226
112346 GCGAAGAGTGGAAAAACAGTGTACAGCTTTTATTTTGGGGGGCTCCAAAATCACTGCAG 112287
2227 ATGTGACTGCAGCCATGAAATTAAGACACTTACTCTCTTGGAGAAAAGTTA-ACCA 2284
112286 ATGTGACTGCAGCCATGAAATTAAGACACTTACTCTCTTGGAGAAAAGTTAATGACCA 112227
2285 ACCTAGATAGCATATGGAAGAGAGAGACATTTACTCTGCCAAACAAAGCCCATCTAGTCA 2344
112226 ACCTAGATAGCATATTTCAATGCGAAGAAACATTTCTTGGCAACAAAGTTGCGACTAGTCA 112167
2345 AGGCTATGCTTTTCCAGTGTGATGTATGGATGAGAGTTGAGCTGTGAAGAAAGCTG 2404
112166 AGGCTATGCTTTTCCAGTGTGATGTATGGATGAGAGTTGAGCTGTGAAGAAAGCTG 112107
2405 AGCACTGAAGAAATGATGCTTTTGAACGTGGTGTGGAGAGACTCTTGAAGTCCCTT 2464
112106 AGCGCAAGAAATGATGCTTTTGAACGTGGTGTGGAGAGACTCTTGGGTGTCCCTT 112047
2465 GGACTGCAAGAGATCCCAACAGATCCATTTCTGAAGAGAGATCAGCCCTGGGATTTCTTTGG 2524
112046 GGACTGCAAGAGATCCCAACAGATCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTGG 111987
2525 AAGGAATGATGCTAAAGCTGAAAACCTCCAGTACTTTTGGCCACCTCATGGAAGAGTGA 2584
111986 AAGGAATGATGCTAAAGCTGAAAACCTCCAGTACTTTTGGCCACCTCATGGAAGAGTGA 111927
2585 CACTGGAAGAGACTGCTGGAGGAAATGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2644
111926 CAATTGGAAGAGACTCTGATGCTGGAGGAAATGGGGGAGAGAGAGAGAGAGAGAGAGAGAG 111867
2645 GGATGAGATGGCTGATGGCATCACTGATCGATCGATCGATGGGTGAATCTCTGG 2704
111866 GGATGAGATGGCTGATGGCATCACTGATCAATGGAAGTGAAGTGGGTGAATCTCTGG 111807
2705 AGTTGGTGAAG 2764
111806 AGTTGGTGAAG 111747

repeat_region	/rpt family="Bov-B LINE/BovB"	22768..22793	Db	223920	TGTGCTTTGCTGGAGAGCGCTGAAGAGAAACCCACACCCCAAGGTAAGAGAAACCCAAAG	223861
repeat_region	/note="(AACTG)n Simple repeat"	complement(24050..24326)	Qy	121	TAAGATGGTAGGTGTTGTGAGAGGGCATCAGAGGGCAGACATCTGAAACCATACAGGCA	180
repeat_region	/rpt family="Bov-A2 SINE/BovA"	24649..24847	Db	223860	TAAGATGGTAGGTGTTGCAAGAGGGCATCAGAGGGCAGACACACTGTAACCATCTACAC	223801
repeat_region	/rpt family="L1MDA LINE/L1"	24879..25119	Qy	181	GAAACTAGTCAATCAATCACAACACTAGGACCAAGCCCTGCTTAACCTCAATGAACCTAAG	240
repeat_region	/rpt family="L1MDA LINE/L1"	25113..25494	Db	223800	GAAACATAGTCAATCTATTACACTAGGACCAAGCCCTGCTTAACCTCAATGAATTAAG	223741
repeat_region	/rpt family="L1MCI LINE/L1"	25638..25736	Qy	241	CCATGCCCTGGGGCAACCCCAAGATGGCGGGTCACTGGTGGAGAGATCTGCACAGAAATGTG	300
repeat_region	/rpt family="Bov-D SINE/BovD"	25795..25989	Db	223740	CCATGCCCATGGGGCAACCCCAAGATGGCGGGTCACTGGTGGAGAGATCTGCACAGAAATGTG	223681
repeat_region	/rpt family="L1MDA LINE/L1"	complement(26231..26231)	Qy	301	GTCACATGGAGAGAGGAAT-GCAAAACCACTTCAGTATCTTGTCTTCTGAGAACCCCATGAA	359
repeat_region	/rpt family="L1MCI LINE/L1"	complement(26237..26538)	Db	223680	GTCACATGGAGAGAGGAATGGCAACCACTTCAGTATCTTGTCTTCTGAGAACCCCATGAA	223621
repeat_region	/rpt family="Bov-B LINE/BovB"	26545..26564	Qy	360	CAGTATGAAAAGGCAAAATGATAGTACTGAAAGAGGAACTCCCCAGGTCACTAGGTGC	419
repeat_region	/note="(CAGAT)n Simple repeat"	complement(26565..26875)	Db	223620	CAGTATGAAAAGGCAAAATGATAGTACTGAAAGAGGAACTCCCCAGGTCACTAGGTGC	223561
repeat_region	/rpt family="Bov-B LINE/BovB"	complement(26872..27574)	Qy	420	CCCATACTGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAGCC	479
repeat_region	/rpt family="L1MCI LINE/L1"	complement(27620..27740)	Db	223560	TCAATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAGCC	223501
repeat_region	/rpt family="Bov-D SINE/BovD"	complement(27741..27863)	Qy	480	AAACCAAAAGAAATACCCAGCTGGATGTGACTGGTGATATAAGCAAGGTCCCATGCTG	539
repeat_region	/rpt family="L1MCI LINE/L1"	28254..28374	Db	223500	AAAGCAAAAGAAATACCCAGATGTGGATGTGACTGGTGATATAAGCAAGGTCCCATGCTG	223441
repeat_region	/rpt family="Bov-ta1 SINE/BovA"	28464..28568	Qy	540	TAAAGAGCAATATTGCAATAGGAACTGCAATGTGAGTCCATGATTAAGCAAGGCAAAATTTGA	599
repeat_region	/rpt family="Tigger2 DNA/MER2_type"	complement(28575..28751)	Db	223440	TAAAGAGCAATATTGCAATAGGAACTGCAATGTGAGTCCATGATTAAGCAAGGCAAAATTTGA	223381
repeat_region	/rpt family="Bov-A2 SINE/BovA"	28766..28858	Qy	600	AGTGGTCAAAACAGAGATGGCAAGTGCAATGTCAACATCTAGGAATCAGCGAACTAAA	659
repeat_region	/rpt family="Tigger2 DNA/MER2_type"	28860..29023	Db	223380	AGTGGTCAAAACAGAGATGGCAAGTGCAATGTCAACATCTAGGAATCAGCGAACTAAA	223321
repeat_region	/rpt family="L1MB8 LINE/L1"	29030..29105	Qy	660	ATGGACTGGAAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTCGGGCAGGA	719
repeat_region	/rpt family="Bov-D SINE/BovD"	29106..29126	Db	223320	ATGGACTGGAAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTCGGGCAGGA	223261
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	30407..30840	Qy	720	ATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCCGAAATGCACTACT	779
repeat_region	/rpt family="L1MB8 LINE/L1"	30841..31025	Db	223260	ATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCCGAAATGCACTACT	223201
repeat_region	/note="(CAGCT)n Simple repeat"	31026..31050	Qy	780	TGGATGCAATCTCAAAAACGACAGATGATCTCTGTTTGTTCCTCAAGGCAAAACCATTCAA	839
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	31051..31241	Db	223200	TGGATGCAATCTCAAAAACGACAGATGATCTCTGTTTGTTCCTCAAGGCAAAACCATTCAA	223141
repeat_region	/rpt family="L1MDA LINE/L1"	31242..31296	Qy	840	TATCA CAGTAATCCAAAGTCTATGCCCCAACCAAGTAATGCTGCAAGAGCTGAAGTTGAACG	899
repeat_region	/rpt family="Tigger2 DNA/MER2_type"	complement(31320..31509)	Db	223140	TATCA CAGTAATCCAAAGTCTATGCCCCAACCAAGTAATGCTGCAAGAGCTGAAGTTGAACG	223081
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	complement(31510..31625)	Qy	900	GTCCTATGAAGACCTCAGAGACCTTTTAGAACTAACCCCAAAAAGATGTCCTTCTCAT	959
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	31626..31897	Db	223080	GTCCTATGAAGACCTCAGAGACCTTTTAGAACTAACCCCAAAAAGATGTCCTTCTCAT	223021
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	31898..31921	Qy	960	TATAGGGAGCTGGAATGCAAAAGTAGGAAGCAAAACACCTCGAGTAACACAGGCAAAAT	1019
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	31922..31997	Db	223020	TATAGGGAGCTGGAATGCAAAAGTAGGAAGCAAAACACCTCGAGTAACACAGGCAAAAT	222961
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	31998..32021	Qy	1020	TGGCTTGGAAATACCGAATGAAGAGGGGCAAAAGACTAATAGAGTTTTCGCAAGAAATGC	1079
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	32022..32097	Db	222960	TGGCTTGGAAATACCGAATGAAGAGGGGCAAAAGACTAATAGAGTTTTCGCAAGAAATGC	222901
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	32098..32121	Qy	1080	ACTGGTTCATGCAAAACACCC--TCTTCCAAACCAACAGAGAGAAAGACTCTATACATGGACA	1137
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	32122..32197	Db	222900	ACTGGTTCATGCAAAACACCC--TCTTCCAAACCAACAGAGAGAAAGACTCTATACATGGACA	222841
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	32198..32271	Qy	1138	TCAACAGATGGTCAACACCCAAATCAGATTGATTATTTCTTTTGCAGCAAGATGGAGA	1197
repeat_region	/rpt family="Bov-ta2 SINE/BovA"	32272..32347	Db	222840	TCAACAGATGGTCAACACCTGAAATCAGATTGATTATTTCTTTTGCAGCAAGATGGAGA	222781

Query Match 60.48; Score 2158.67; DB 1; Length 233345;
 Best Local Similarity 72.08; Pred. No. 0;
 Matches 2910; Conservative 0; Mismatches 468; Indels 561; Gaps 14;
 1 GTCAGGAGGGCGGCGAGTGGAGAGGAGTACTACCTCGTCCAAAGGTAAGAGAGTACG 60
 223975 GTCAGGAGGGCGGCGTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 223921
 61 TGGCTTTGCTGGAGCAGCCGTAAGAGATACCCACGCGCCCAAGGTAAGAGAAACCCAAAG 120

[illegible]

[illegible][illegible]

TITLE
JOURNAL REFERENCE
AUTHORS TITLE
JOURNAL REFERENCE
AUTHORS TITLE
JOURNAL COMMENT

NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 198916)
Green, E.D.
Direct Submission
Submitted (30-MAY-2003) NIH Intramural
Grovmont Circle, Gaithersburg, MD 20885
3 (bases 1 to 198916)
Green, E.D.
Direct Submission
Submitted (03-OCT-2003) NIH Intramural
Grovmont Circle, Gaithersburg, MD 20885
On Oct 3, 2003 this sequence version 1
----- Genome Center
Center: NIH Intramural Sequencing
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc_zoo@nigr1.nih.gov
----- Project Information
Center project name: eeo
Center clone name: 385L09

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.950319
Consensus quality: 197481 bases at least Q40
Consensus quality: 197872 bases at least Q30
Consensus quality: 198094 bases at least Q20
Insert size: 153000; agarose-fp
Insert size: 198216; sum-of-contents
Quality coverage: 15.52% in Q20 bases; agarose-fp
Quality coverage: 11.98% in Q20 bases; sum-of-contents

* NOTE: This is a 'working draft' sequence. It currently

* consists of 8 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

*	1	19088:	contig of 19088 bp in length
*	19089	19188:	gap of unknown length
*	19189	21334:	contig of 2146 bp in length
*	21335	21434:	gap of unknown length
*	21435	37376:	contig of 15942 bp in length
*	37377	37476:	gap of unknown length
*	37477	109937:	contig of 72461 bp in length
*	109938	110037:	gap of unknown length
*	110038	157979:	contig of 47942 bp in length
*	157980	158079:	gap of unknown length
*	158080	160129:	contig of 2050 bp in length
*	160130	160229:	gap of unknown length
*	160230	170532:	contig of 10303 bp in length
*	170533	170632:	gap of unknown length
*	170634	198916:	contig of 28284 bp in length

```

FEATURES
  Location/Qualifiers
    1. .198916
      /organism="Bos taurus"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-38519"
/clone_lib="RP42"
1..23685
/note="clone overlaps with GenBank Accession Number
AC14802 clone RP42-212N23 (center project name een)"
misc_feature
1..19088
/note="assembly_fragment
clone_end:77
vector_side:left"
19189..21334
/note="assembly_fragment"
21435..37376
/note="assembly_fragment"
37477..109937
/note="assembly_fragment"
110038..157979
/note="assembly_fragment"
158080..160129
/note="assembly_fragment"
160230..170532
/note="assembly_fragment"
170633..198916
/note="assembly_fragment
clone_end:SP6
vector_side:right"

Query Match      60.3%; Score 2153.4; DB 1; Length 198916;
Best Local Similarity 77.5%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 156; Indels 606; Gaps 8;

QY 1 GTCAGGAGCGCGCAGTGGAGAGTACCTACCTCTGTCGCAAGGTAAAGGACGACGTAGC 60
Db 198496 GTCAGGAGAGCGGTGTGTAGGAGA-----TAACCTCTGTCGCAAGGTAAAGGACGACGTAGC 198442
QY 61 TGCCTTTTGTGGAGCGCGTAAAGAGATACCCCGCCCAAGGTAAAGAGAACCCCAAG 120
Db 198441 TGTCTTTTGTGGAGCGCGGTGAAGAGATATCCAGCTCCAGGTAAAGAGAACCCCAAG 198382
QY 121 TAAGATGGTAGTGTGTGGAGGGCATCAGAGGGCAGACATCTGAACCATACACGCA 180
Db 198381 TAAATGGTAGTGTGTGAAGGGCATCAGAGGGCAGACATCTGAACCATACACGCA 198322
QY 181 GAAACTAGTCAATCTAATCACACTAGGACACAGCCCTGTCTTAACCTCAATGAACCTAAG 240
Db 198321 GAAACTAGTCAATCTAATCACACTAGGACACAGCCCTGTCTTAACCTCAATGAACCTAAG 198262
QY 241 CCATGCCGTGGGCAACCCCAAGATGGCGAGGTCTGTTGGAGAGATCTGACAGAAATGTG 300
Db 198261 CCATGCCGTGGGCAACCCCAAGATGGCGAGGTCTGTTGGAGAGATCTGACAGAAATGTG 198202
QY 301 GTCCACTGGAGAGAGGGAAT-GCAACCACTTCAGTATCTCTGCTTGAGAACCCCATGAA 359
Db 198201 GTCCACTGGAGAGAGGGAATGCAACCACTTCAGTATCTCTGCTTGAGAACCCCATGAA 198142
QY 360 CAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACTCCCGAGGTCAAGTGGTC 419
Db 198141 CAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACTCCCGAGGTCAAGTGGTC 198082
QY 420 CCATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGAGATGGAGCC 479
Db 198081 CCAATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGAGATGGAGCC 198022
QY 480 AAAGCAAAAAGAAATACCCAGCTGTGGATGTGACTGTGTGATATAAGCAAGGTCCGATGCTG 539
Db 198021 AAAGCAAAAAGAAATACCCAGCTGTGGATGTGACTGTGTGATACAGCAAGGTCCGATGCTG 197962
QY 540 TAAAGAGCATATTTGATAGGAACCTTGGAAATGTGAGGTCCATGAATCAAGGCAAAATGGA 599
Db 197961 TAAAGAGCATATTTGATAGGAACCTTGGAAATGTGAGGTCCATGAATCAAGGCAAAATGGT 197902
QY 600 AGTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATCTTAGGAAATCAGGAACTAAA 659

```

```

Db 197901 AGTGGTCAAAACAGAGATGGCAAGAGTGAATCTCGACATCTTAGGAATCAGGAACTGAA 197842
QY 660 ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTAATCTACTACTACTGCGGCAGGA 719
Db 197841 ATGGACTGGAATGGGTGAATTTAACTGAGATGACCATTAATCTACTACTACTGCGGCAGGA 197782
QY 720 ATCCCTCAGAGAAATGGAGTAGCCATCATGTCTCAACAAAAGAGTCCGAAATTCAGTACT 779
Db 197781 ATCCCTCAGAGAAATGGAGTGGCCATCATGTCTCAACAAAAGAGTCTGAATTCAGTACT 197722
QY 780 TGGATGCACTCTCAAAAACGACAGAAATGATCTCTGTTTGTTCAGAGCAAAACCAATCAA 839
Db 197721 TGGATGCACTCTCAAAAACGACAGAAATGATCTCTGTTTGTTCAGAGCAAAACCAATCAA 197662
QY 840 TATCACAGTAATCCCAAGTCTATGCCCCCAACCAAGTAATCTGTAAGAGCTGAAGTTGAACG 899
Db 197661 TATCACAGTAATCCCAAGTCTATGCCCCCAACCAAGTAATCTGTAAGAGCTGAAGTTGAACG 197602
QY 900 GTCCTATGAAGACCTTACAGACCTTTTAGAACTAACACCCCAAAAAGAGTGTCTTCTCAT 959
Db 197601 GTCCTATGAAGACCTTACAGACCTTTTAGAACTAACACCCCAAAAAGAGTGTCTTCTCAT 197542
QY 960 TATAGGGACCTGGAAATGCAAAAAGTAGGAAGCAAGAAACACCTGGAGTAACAGGCAATTT 1019
Db 197541 TATAGGGACCTGGAAATGCAAAAAGTAGGAAGTAAGAAACACCTGGAGTAACAGGCAATTT 197482
QY 1020 TGGCCTTGGAAATACGGAATGAAGCAGGCAAAAGACTAATAGAGTTTTCGCAAGAAATGC 1079
Db 197481 TGGCCTTGGAAATACGGAATGAAGCAGGCAAAAGACTAATAGAGTTTTCGCAAGAAATGC 197422
QY 1080 ACTGTCTATGCAAAACACCCCTCTTCCAAACAACAAGAAAGACTCTACATGGACATC 1139
Db 197421 ACTGTCTATGCAAAACACCCCTCTTCCAAACAACAAGAAAGACTCTACATGGACATC 197362
QY 1140 ACCAGATGGTCAACACCGAAATCAGATTGATTATTTCTTTGCGACCAAAAGATGGAGAAG 1199
Db 197361 ACCAGATGGTCAACACCGAAATCAGATTGATTATTTCTTTGCGACCAAAAGATGGAGAAG 197302
QY 1200 CTCTATACAGTCAGCAAAAACCAAGACCAAGGAGCTTACTTGGCTCAGATCATGAATCTCT 1259
Db 197301 CTCTATACAGTCAGCAAAAACCAAGACCGGAGCTGACTTGGCTCAGATCATGAATCTCT 197242
QY 1260 TATTGCCAAATTCAGACTTAAATTTGAAGAGTAGGGAACCACTAGACTACCTCAGCTA 1319
Db 197241 TATTGCCAAATTCAGACTTAAATTTGAAGAGTAGGGAACCACTAGACTACCTCAGCTA 197182
QY 1320 AGACTTAAATCCATCCCTTATATGATTATACAGTGGAAAGTGAAGAAATAGATTAAAGGCTCT 1379
Db 197181 GGACTTAAATCCATCCCTTATGATTATACAGTGGAAAGTGAAGAAATAGATTAAAGGCTCT 197122
QY 1380 AGATCTGATAGACAGAGTACCTAATGAACCTATGCAAGAGGTTCTACATTTGACAGGA 1439
Db 197121 AGATCTGATAGAGTGGCTGATGAACCTATGCAAGAGGTTTGTGACATTTGACAGGA 197062
QY 1440 CACAGGATCAGAGACCATCCCATGGAAAGAAATGCAAAAAGCAAAATGGCTGTCTGG 1499
Db 197061 CACAGGATCAGAGACCAT-CCCATGGAAAGAAATGCAAAAAGCAAAATGGCTGTCTGG 197003
QY 1500 GGAGGCTCTTCAAAATAGCTGTGAAAAGAGAGAGTGAAGCAAGGAAAAAGGAAAG 1559
Db 197002 GGAGGCTCTTCAAAATAGCTGTGAAAAGAGAGAGTGAAGCAAGGAAAAAGGAAAG 196943
QY 1560 ATAAAGCATCTGAATCCAGAGTTCCAAA----- 1588
Db 196942 ATATAACATCTGAATCCAGAGTTCCAAAATAGCAGAGAGATGAAGAAAGCTTCTT 196883
QY 1589 ----- 1588
Db 196882 CAGCAAAATGCAAGAAATAGAGGAAAAACAAGAAATGGGAAAAAGACTAGGATCTCTT 196823
QY 1589 ----- 1588

```


Db 196822 CAAGAAATAGAGTACCAAGGACATTTTCATGCAAGATGAGCTCGATTAAGGACAG 196763
Qy 1589 ----- 1588
Db 196762 AAATGGTATGACCTAACAGAGCAGAGATATTAAAGAGAGGTGGCAAGAATACACAGA 196703
Qy 1589 ----- 1588
Db 196702 AGAACTGTACAAAAGATCTTCACGACCCAGATATCAGATGGTGTGATCACTGACCT 196643
Qy 1589 ----- 1588
Db 196642 AGAGCCAGACATCCTGGAAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTATGAAGAA 196583
Qy 1589 ----- 1588
Db 196582 AGCTAGTAGAGTATGGAATTCAGTTGAGCTATTCCAAATCCTGAAAGATGATGCTGT 196523
Qy 1589 ----- 1588
Db 196522 GAAAGTGTGCACTCAATATGCGAGCAAAATTTGGAATACTCAGCAGTGGCCACAGGACTG 196463
Qy 1589 ----- 1588
Db 196462 GAAAAAGTCACTTTTCATTCCAATCCCAAGAAAGCAATGCAAAAGATGTTCAAACTA 196403
Qy 1589 ----- 1588
Db 196402 CCGCACAATTGCATCTCAACGCTAGTCAAGTAATGCTCAAAATTTCTCCAAAGCCAG 196343
Qy 1589 -----GAACTTCAGTTGTTCAAGCTGGTTTATAGAAAGTCAAG 1626
Db 196342 GCTTCAGCAATATGTGAACCTGTTGAACCTTCCTGATGTTCAAGCTGGTTTATAGAAAGCCAG 196283
Qy 1627 AGGAACAGAGACCAATTCGCAACATCCTCTGATCATGGAAGAAAGCAAGAGAGTTCCA 1686
Db 196282 AGGAACAGAGATCAAAATTCGCAACATCCTCTGATCATCAAAAGAGAGAGAGTTCCA 196223
Qy 1687 GAAAAACATCTATTTCCTGCTTTATGCTATGCAAAAGCCTTTGACTGTGGGGTCACAA 1746
Db 196222 GAAAAACATCTATTTCCTGCTTTATGCTATGCTAAAGCCTTTGACTGTGGATCACAA 196163
Qy 1747 TAACTGTGGAATTTCTGAAGGATGGGAATACAGAGACCACTGACCTGACTCTTGA 1806
Db 196162 TAACTGTGGAATTTCTGAAGAGATGGGAATACAGAGACCACTGACTGCTCTTGA 196103
Qy 1807 AAATTTGTATGCTCAGGAGCAACAGATTTAGAACTGGACATGGAACACAGACTGGTT 1866
Db 196102 AAATTTGTATGCTCAGGAGCAATAGTTAGAACTGGACATGGAACACAGACTGGTT 196043
Qy 1867 CCAAGTAGGAAAGAGATATGTCAGGCTGTATATGTCACCGGCTTTTAACTTCTA 1926
Db 196042 CCAATAGGAAAGAGATTCGTCAGGCTGTATATGTCACCGCTTTTAACTTATC 195983
Qy 1927 TGCAGAG-ACATCATGAGAAAGCTGGGCTGGAGAGAGCAAGCTGGAATCAAGATTGC 1985
Db 195982 TGCAGAGTACATCATAGAAAGCGCTGGACTGGAGAGAAACCAAGCTGGAATCAAGATTGC 195923
Qy 1986 CGGAGAGAAATAGCAATTAACCTCAGATATGCAAGATGATACCAACCTTATGCGAGAAAGTGA 2045
Db 195922 CGGAGAGAAATATCAATAACCTCAGATATGCAAGATGATACCAACCTTATGCGAGAAAGGGA 195863
Qy 2046 AGAGAACTAAAAGCCTCTTGTATGAGAGTGAAGAGAGAGTGAAGAGTTGGCTTAAA 2105
Db 195862 AGAGGAACTAAAAGCCTCTTGTATGAGAGTGAAGAGTGAAGAGTTGGCTTAAA 195803
Qy 2106 GCTCAACATTCAGAAAAGCAAGATCATGGCATCTGGTCCCATCACTCTCATGGGAAATAGA 2165
Db 195802 GCTCAACATTCAGAAAAGCAAGATCATGGCATCTGGTCCCATCACTCTCATGGGAAATAGA 195743
Qy 2166 TGGGAAACAGTGGAAACAGTGTGAGACTTTATTTTTTTGGGGGGCTCCAAATCACTGCA 2225
Db 195742 TGGGAAACAGTGGAAACAGTGTGAGACTTTATTTTTTTT-GGGGCTCCAAATCACTACA 195685

Qy 2226 GATGGTGACTCAGCCATGAAATTAAGAGACACTTACTCTTGGAGAAAGTTA--ACC 2283
Db 195684 CATGGTGACTCAGCCATGAAATTAAGAGCGCTTACTCTTGGAGAAAGTTATGACC 195625
Qy 2284 AACCTAGATAGCATATTGAAAGCAGAGACATTAACCTTGGCAACAAAGCCCATCTAGTC 2343
Db 195624 AACCTAGATAGTATATTCAAAAGCAGAGACATTAACCTTGGCGACAAAGGTTTGTCTAGTC 195565
Qy 2344 AAGCTATGTTTTCAGTGGTTCATGTATGATGTGAGAGTGTGAGACTGTGAGAAAGCT 2403
Db 195564 AATGCTATGG-TTTTCTGTGGTTCATGTATGATGTGAGAGTGTGAGACTGTGAGAAAGCT 195506
Qy 2404 GAGCAGCTGAAGAAATTGATGCTTTTGAACCTGTGGTGTGGAGAAAGCTTCTTGGAGAGTCCCT 2463
Db 195505 GAGCGCCAAAGAAATTGATGCTTTTGAACCTGTGGTGTGGAGAAAGCTATTGAGAGTCCCT 195446
Qy 2464 TGGACTGCAAGAGATCAACACAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCTTTG 2523
Db 195445 TGGACTGCAAGGAGATCAACACAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCTTTG 195386
Qy 2524 GAAAGAAATGATGCTAAAGCTGAAACTCCAGTACTTTTGGCCACCTGATCAGAAAGAGCTGAC 2583
Db 195385 GAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTCATGCAAGAGTTGAC 195326
Qy 2584 TCAGTGGAAAAAGACCTGATGCTGGAGGAGATTGGGGCAGAGAGAGAGGAGCGACAG 2643
Db 195325 TCATTGGAAAAAGCTCTGATGCTGGAGGAGATTGGGGCAGAGAGAGAGGAGATGACAG 195266
Qy 2644 AGGATGAGATGCTCGATGGATCACTGACTCGATGAGACGTGAGTCTGGGCTCAACTCCTG 2703
Db 195265 AGGATGAGATGCTCGATGGATCACTGACTCGATGAGACGTGAGTCTCAGTGAAGTCTGA 195206
Qy 2704 GAGTTGGTATGAGCAGAGAGGCGCTGCTCGGGGATTCATGGGTCACAAAGAGTTGG 2763
Db 195205 GAGTTGGTATGAGCAGAGAGGCGCTGCTCGGATTCATGGGTCGCAAGAGTTGG 195146
Qy 2764 ACACACTGAGCAACTGAACTGAACTGAACTG 2795
Db 195145 ACACACTGAGCAACTGATCTGATCTGATCTG 195114

RESULT 4
AC135839/c

LOCUS Bos taurus clone RP42-2H17, WORKING DRAFT SEQUENCE, 6 ordered pieces
DEFINITION AC135839
ACCESSION AC135839
VERSION AC135839.2 GI:24942893
KEYWORDS HTG: HTGS_PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 179015)
AUTHORS Akhtar N., Antonellis A., Ayele K., Beckstrom-Sternberg S.M., Benjamin B., Blakesley R.W., Bouffard G.G., Brinkley C., Brooks S., Carls G., Coleman B., Engle J., Granite S., Guan X., Gupta J., Haghghi P., Han J., Hansen N., Ho S.-L., Idol J.R., Karlins E., Laric P., Lee-Lin S.-Q., Legaspi R., Maduro Q.L., Maduro V.B., Margulies E.H., Masello C., Maskeri B., McDowell J., Paquirigan C., Pearson R., Portnoy M.E., Prasad A., Reddix-Dugue N., Schandler K., Schueler M.G., Sison C., Statropoulos S., Thomas J.W., Thomas P.J., Touchman J.W., Vogt J.L., Wetherby K.D., Wiggins L., Young A. and Green E.D.
TITLE NISC Comparative Sequencing Initiative
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 179015)
AUTHORS Green E.D.
DIRECT SUBMISSION
TITLE Submitted (23-OCT-2002) NIH Intramural Sequencing Center, 8717
JOURNAL Grovemont Circle, Gaithersburg, MD 20877, USA

QY	839	ATATCACAGTAATCCAAAGTCTATGCCCCCAACCAAGTAATGCTCAGAAAGCTGAAGTTGAAC	898	Db	144760	GCTAGTGGAGGTGATAGAAATTCAGTTGAGCTATTCCAAATCCTGAAAGATGATGCTGTG	144701
Db	145839	ATATCACAGTAATCCAAAGTCTGCCCCCAACCAAGTAATGCTCAGAAAGCTGAAGTTGAAC	145780	QY	1589	-----	1588
QY	899	GGTCTATGAAGACCTTACAAGACCTTTTGAAGACTAACACCCCAAAAAAAGATGTCCTTCTCA	958	Db	144700	AAAGTGTGCTGACCTCAATATATGCCAGCAAAATTTGGAAAACTCAGCAGTGACCACAGGACTGG	144641
Db	145779	GGTTCTATGAAGACCTTACAAGACCTTTTGAAGACTAACACCCCAAAAAAAGATGTCCTTCTCA	145720	QY	1589	-----	1588
QY	959	TTATAGGGGACTGGAAATGCAAAAGTAGGAAGCAAAAGAAACACCTGAGTAAACAGCAAAAT	1018	Db	144640	AAAAGGTCAAGTTTTCATTCCTCAATCCCAAGAAAGCAATGCCAAAGAAATGCTCAAACTAC	144581
Db	145719	TTATAGGGGACTGGAAATGCAAAAGTAGGAAGCAAAAGAAACACCTGAGTAAACAGCAAAAT	145660	QY	1589	-----	1588
QY	1019	TTGGCCCTTGAATACGGAATGAAGCAGGCAAGACTAATAGATTTTGGCCCAAGAAAAATG	1078	Db	144580	CACACAATTTGCACTCATCTCACATGCTPAATAAAGTAATACTCAAAATTTCTCCAAAGCCAGG	144521
Db	145659	TTGGCCCTTGAATACGGAATGAAGCAGGCAAGACTAATAGATTTTGGCCCAAGAAAAATG	145600	QY	1589	-----	1588
QY	1079	CACCTGTATAGCAAAACACCCCTTCTCCAAACAACAAGAAAGACTCTACACATGGACAT	1138	Db	144520	CTTCAGCAATACGTGAACCGTGAACCTTCAGATGTTCCAGATGTTTGAAGAAAGGCAGA	144461
Db	145599	CACCTGTATAGCAAAACACCCCTTCTCCAAACAACAAGAAAGACTCTACACATGGACAT	145540	QY	1628	GGAAACCAAGAGACCAAAATTTGCCAACATCCTCTGTATCATGGAATAAGCAAGAGAGTTCCAG	1687
QY	1139	CACCAAGTGTCAACACCGAAATCAGATTGATTATATCTTTTGCAGCCCAAGATGGAGAA	1198	Db	144460	GGGACCGAGAGACCAAAATTTGCCAACATCCTGCTGGATCATGGAATAAGCAAGAGAGTTCCAG	144401
Db	145539	CACCAAGTGTCAACACCGAAATCAGATTGATTATATCTTTTGCAGCCCAAGATGGAGAA	145480	QY	1898	AAAAACATCTATTTCTGCTTTTATGACTATGCAAAAGCCTTTTGACTGTGGGGGTCACAAT	1747
QY	1199	GCTCTATACAGTACAGCAAAAAACAAGACCAAGAGCTTTACTGTGGCTCAGATCATGAATCC	1258	Db	144400	AAAAGCATCTTATTTCTGCTTTTATGCTTTTGGCGATGCCAAAGCCTTTAAGTGTGGGTCAAT	144341
Db	145479	GCTCTATACAGTCAAC-CAACAAGACGGGAGCTGACTGTGGCTCAGATCATGAATCC	145421	QY	1748	AACTGTGGAAAAATTTCTGAAAGGATGGGAATACCAGACCACTGACCTCTTGAAA	1807
QY	1259	TTATTTGCCAAAATTCAGACTTTAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGT	1318	Db	144340	AACTGTGGACGATTTCTGAAAGACATGGGAATACCAGACCACTGACCTCTCTTGAGA	144281
Db	145420	TTATTTGCCAAAATTCAGACTTTAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGT	145361	QY	1808	AATTTGTATGAGTTCAGGAAGCAACAGTTAGACTGGACATGGAACAACAAGACTGGTTC	1867
QY	1319	AAGACCTTAATCCAAATCCCTTATGATTATACAGTGGAAAGTAGAAATAGATTAAAGGGCC	1378	Db	144280	AATCTGTATGAGGTTCAGGAAGCAACAGTTAGACTGGACATGGAACAACAAGACTGGTTC	144221
Db	145360	ATGACCTTAATCCAAATCCCTTATGATTATACAGTGGAAAGTAGAAATAGATTAAAGGGCC	145301	QY	1868	CAAGTAGGAAAAAGAGTAGTTCAGAGGTGTATATTGTTCACCGGCTTGTTTAACTTCTAT	1927
QY	1379	TAGATCTGATAGACAGAGTACCTTAATGAATATGACAGAGTTTCAGATCATGTTACAGG	1438	Db	144220	CAATAGACAGAGGATGACTCAAGGCTGTATATTGTTCACCTGCTTATTAACTTATAT	144161
Db	145300	TACATCTGATAGATAGAGTGCCTTCATGAGTATGGAATGAGGTTTGTGACACTGACAG	145241	QY	1928	GCAGA-GACATCATAGAAAAAGCTGGGCTGGAAGACACAGCTGGATCAAGATTGCC	1986
QY	1439	AGACGGGATCGAGACCATCCCATGGAAGAAAGTGCAGAAAGCAAAATGCTGCTCTG	1498	Db	144160	GCAGATTACATCATGAGAAAAAGCTGGAAGCAACCAAGCTGGAATCAAGATTGCC	144101
Db	145240	AGACGGGATCGAGACCATCCCATGGAAGAAAGTGCAGAAAGCAAAATGCTGCTCTG	145181	QY	1987	GGGAGAAATAGCAATAACCTCAGATATGCAGATATACCACTTATGGCAAGAAAGTGA	2046
QY	1499	GGGAGGCTTACAATAGCTGTGAAAGAGAGAGAGTGAAGCAAGCAAGAAAAAGGAAA	1558	Db	144100	GGGAGAAATATCATTAACCTCAGATATGCAGATGACACCACTTATGGCAAGAAAGTGA	144041
Db	145180	GGGAGGCTTACAATAGCTGTGAAAGAGAGAGAGTGAAGCAAGCAAGAAAAAGGAAA	145121	QY	2047	GAGGAACTAAAAAGCCTCTTGATGAAGGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAG	2106
QY	1559	GATAAAAGCATCTGAATGCAAGTTCCAAA-----	1588	Db	144040	GAGGAACTAAAAAGCCTTTTGATGAAGGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAG	143981
Db	145120	GATATAAGCATGTGAATGCAAGTTCCAAAAGATAGCAAGAGAGATAAGAAAGCCTTCC	145061	QY	2107	CTCAACATTCAGAAAAAGCAAGATCATGGCATCTGGTCCCATCATCTCATGGGAAATAGAT	2166
QY	1589	-----	1588	Db	143980	CTCGAATTCAGAAAAAGCAAGATCATGGCATCCGGTCCCATCATCTCATGGGAAATAGAT	143921
Db	145060	TCAGCGATCAATGCAAGAAATAGAGGAAACCAACAGAAATGGTAAGACTAGAGATCTCT	145001	QY	2167	GGGGAAAAAGCTGGAACAGTGTGACACTTTATTTTTGGGGGGCTCCAAAATCATCTGAG	2226
QY	1589	-----	1588	Db	143920	GGGGAAAAAGCTGGAAGAGAGTGGCTGACTTTA--TTTTGGGGGGCTCCAAAATCATCTGAG	143863
Db	145000	TCAGAAAAATTAGAGATACCAAGGGGACATTTTCATGCAAGAGATGGGCTTGATAAGGGCA	144941	QY	2227	ATGGTGAAGTGCAGCCATGAAATTAAGACACTTACTCTCTTGGGAAAGAAAGTTA--ACCA	2284
QY	1589	-----	1588	Db	143862	ATGGTGAAGTGCAGCCATGAAATTAAGACACTTACTCTCTTGGGAAAGAAAGTTAAGCA	143803
Db	144940	GAAATGTAGGGACCTTAACAGAACAGAGATATTAAGAAAGAGATGGCAAGATACACAG	144881	QY	2285	ACCTAGATAGCATATTGAAAGACAGAGATTTACCTTGCACCAAGAGGCTTCTAGTCA	2344
QY	1589	-----	1588	Db	143802	AGCTAGATAGCATATTGAAAGACAGAGATTTACTTTTGGGAAACAAAGGCTCTCTAGTCA	143743
Db	144880	AAGAACTGTACAAAAAGATCTTTCACGACCCAGGATAATCAGATGGTGTGATCACTCATC	144821	QY	2345	AGGCTATGTTTTTCCAGTGGTCACTGTATGGATGTGAGAGTTGGACTGTGGAAGAAAGCTG	2404
QY	1589	-----	1588	Db	143742	AGGCTATGTTTTTCTGTGGTCACTGTATGGATGTGAGAGTTGGACTGTGGAAGAAAGCTG	143683
Db	144820	TAGAGCAGACATCTGGAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTACAAACAA	144761	QY	2405	AGCCTGAAGAAATGATGCTTTTGAAGTGTGGTGTGGAGAGACTCTTTGAGAGTCCCTT	2464
QY	1589	-----	1588				

```

Db      143682 AGCGCTGAAGATTCATGCTTTTGAACCTGTGGTGTGGAGAAGACTCTTAAGAGTCCCTT 143623
QY      2465  GGACTGCAAGAGATCCCAACAGATCCATCTGAGAGGATCAGCCCTGGGATTTCTTTGG 2524
Db      143622 GGACTGCAAGAGATCCCAACAGATCCATCTGAGAGGATCAGCCCTGGGATTTCTTTGG 143563
QY      2525  AAGGAATGATCTTAAGCTGAACTCCAGTACTTTGGCCACTGATCAGAAAGAGCTGACT 2584
Db      143562 AAGGAATGATCTGAAGCTGAACTCCAGTACTTTGGCCACTGATCAGAAAGAGTCTGACT 143503
QY      2585  CACTGGAAGACCTGATGCTGGAGGAGTGGGGCAGAGAGAGAGGGGACGACAGA 2644
Db      143502 CATGGAAAGAACTCTGATGCTGGAGGAGTGGGGCAGAGAGAGGGGACGACAGA 143443
QY      2645  GGATGAGATGGCTGGATGGCATCACTGACTCGATGGAGCTGAGTCTGGGTGAACCTCTCTGG 2704
Db      143442 GGATGAGTGGCTGGATGGCATCACTGACTCGATGGAGCTGAGTCTGGGTGAACCTCTCTGG 143383
QY      2705  AGTTGGTATGACAGAGGAGCCCTGCTCTGGCGGATTCATGGGTCACAAAAGATTGGA 2764
Db      143382 AGTTGGTATGACAGAGGAGCCCTGCTCTGGCGGATTCATGGGTCACAAAAGATTGGA 143323
QY      2765  CACGACTGAGCAACTGAACTGAACTGAACGTGA 2797
Db      143322 CACGACTGAGCAACTGAACTGAACTGAACGTGA 143290

RESULT 5
AC137656/c
LOCUS      AC137656      241923 bp      DNA      linear      HTG 11-MAR-2003
DEFINITION Bos taurus clone RP42-75B10, WORKING DRAFT SEQUENCE, 13 ordered
            pieces.
ACCESSION  AC137656
VERSION    AC137656.3  GI:28913277
KEYWORDS   HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE     Bos taurus (cow)
ORGANISM   Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
Axther,N., Antonellis,A., Avele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Carriaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
Unpublished
NISC Comparative Sequencing Initiative
2 (bases 1 to 241923)
Green,E.D.
Direct Submission
Submitted (11-MAR-2003) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 241923)
Green,E.D.
Direct Submission
Submitted (27-NOV-2002) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
On Mar 11, 2003 this sequence version replaced gi:27764650.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: crb
Center clone name: 075B10

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 239390 bases at least Q40
 Consensus quality: 240096 bases at least Q30
 Consensus quality: 240547 bases at least Q20
 Insert size: 187000; agarose-fp
 Insert size: 240723; sum-of-contigs
 Quality coverage: 13.42x in Q20 bases; agarose-fp
 Quality coverage: 10.42x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 13063: contig of 13063 bp in length
 * 13064 13163: gap of unknown length
 * 13164 20877: contig of 7714 bp in length
 * 20878 20977: gap of unknown length
 * 20978 30802: contig of 9825 bp in length
 * 30803 30902: gap of unknown length
 * 30903 55941: contig of 25039 bp in length
 * 55942 56041: gap of unknown length
 * 56042 86898: contig of 30857 bp in length
 * 86899 86998: gap of unknown length
 * 86999 123280: contig of 36282 bp in length
 * 123281 123380: gap of unknown length
 * 123381 158900: contig of 35520 bp in length
 * 158901 159000: gap of unknown length
 * 159001 171357: contig of 12357 bp in length
 * 171358 171457: gap of unknown length
 * 171458 178118: contig of 6661 bp in length
 * 178119 178218: gap of unknown length
 * 178219 183713: contig of 5495 bp in length
 * 183714 231843: gap of unknown length
 * 231844 231942: contig of 48029 bp in length
 * 231943 237141: gap of unknown length
 * 237142 237241: contig of 5199 bp in length
 * 237242 241923: gap of unknown length
 * 241923 241923: contig of 4682 bp in length.

FEATURES

Location/Qualifiers
 1. .241923
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-75B10"
 /clone_lib="RP42"
 1. .97566
 /notes="clone overlaps with GenBank Accession Number AC135839 clone RP42-2H17 (center project name cra)"
 1. .13063
 /notes="assembly_fragment
 clone_end:SP6
 vector_side:left"
 13164. .20877
 misc_feature
 misc_feature
 misc_feature

misc_feature	/note="assembly_fragment"	Qy	599	AAGTGGTCAAAACAGAGATGCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAATAA	658
misc_feature	20978..30802	Db	66249	AAAGTGGTCAAAACAGAGATGCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAATAA	66190
misc_feature	/note="assembly_fragment"	Qy	659	AATGGACTGGAATGGTGAATTTAACTCAGATGACCAATATATATCTACTACGCGGAGG	718
misc_feature	56042..86898	Db	66189	AATGGACTGGAATGGTGAATTTAACTCAGATGACCAATATATATCTACTACGCGGAGG	66130
misc_feature	/note="assembly_fragment"	Qy	719	AATCCCTCAGAAGAAATGGAGTAGCCATCATGTGTCAACAAAGAGTCCGAATGCAAGTAC	778
misc_feature	12381..158900	Db	66129	AATCCCTCAGAAGAAATGGAGTAGCCATCATGTGTCAACAAAGAGTCCGAATGCAAGTAC	66073
misc_feature	/note="assembly_fragment"	Qy	779	TTGGATGAGTCTCAAAACACAGATGATCTCTGTGTTGTTTCCAGGCGCAACCATTCAC	838
misc_feature	171458..178118	Db	66072	TTGGATGCAATCTCAAAATGACAGATGATCTCTGTGTTTCCAGGCGCAACCATTCAC	66013
misc_feature	/note="assembly_fragment"	Qy	839	ATATCACAGTAAATCCAGTCTGTGCCCAACAGTAAATGCTGAAGAGCTGAGTTGAAC	898
misc_feature	183814..231842	Db	66012	ATATCACAGTAAATCCAGTCTGTGCCCAACAGTAAATGCTGAAGAGCTGAGTTGAAC	65953
misc_feature	/note="assembly_fragment"	Qy	899	GGTCTATGAAGACTCAAGACCTTTTGAAGTAAACACCCAAAGAGATGTCCTTCTCA	958
misc_feature	231943..237141	Db	65952	GGTCTATGAAGACTCAAGACCTTTTGAAGTAAACACCCAAAGAGATGTCCTTCTCA	65893
misc_feature	/note="assembly_fragment"	Qy	959	TTATAGGGGACTGGNAATGCAAGTAGCAAGCAAGAAACACCTGGAGTAAACAGCAAT	1018
misc_feature	237289..241923	Db	65892	TTATAGGGGACTGGNAATGCAAGTAGCAAGCAAGAAACACCTGGAGTAAACAGCAAT	65833
misc_feature	/note="clone overlaps with GenBank Accession Number AC136880 clone RP42-70M11 (center project name crd)"	Qy	1019	TTGGCTTTGGAAATACGGAAATGAAGCAGGGCAAGAACTAATAGAGTTTTCACCAAGAAATG	1078
misc_feature	60.1%; Score 2147.7; DB 1; Length 241923;	Db	65832	TTGGCTTTGGAAATACGGAAATGAAGCAGGGCAAGAACTAATAGAGTTTTCACCAAGAAATG	65773
misc_feature	77.6%; Pred. No. 0; Indels 607; Gaps 9;	Qy	1079	CACTGGTATAGCAAAACCCCTTTCACCAACACACAGAGAGACTCTACACATGGACAT	1138
misc_feature	Matches 2633; Conservative 0; Mismatches 153; Indels 607; Gaps 9;	Db	65772	CACTGGTATAGCAAAACCCCTTTCACCAACACACAGAGAGACTCTACACATGGACAT	65713
1	GTGAGGAAGGCGCGAGTGAAGAGGTACTTACCCTGTCCTCAAGTAAGGACAGTAGC	Qy	1139	CACCAAGTGTCAACACCGAAATCAGATTGATTAATTTCTTGACGCCAAAGATGGAGAA	1198
66844	GTGAGGAAGGCGCGGTGAGGAGA-----TACCCTGTCCTCAAGTAAGGACAGCGGC	Db	65712	CACCAAGTGTCAACACCGAAATCAGATTGATTAATTTCTTGACGCCAAAGATGGAGAA	65653
61	TGGCTTTGCTGAGCGCCGTAAGAGATACCCAGCGCCACAGG-TAAGAGAAACCAA	Qy	1199	GCTCTATACGTACGCAAAACCAAGACCAAGAGCTTACTGTGGCTCAGATCATGAACCTCC	1258
66789	TGGCTTTGCTGAGCGCCGTAAGAGATACCCAGCGCCACAGGTTAAGAGAAACCAA	Db	65652	GCTCTATACGTACCAAC-CAAAACAGACGGGAGCTGACTGTGGCTCAGATCATGAACCTCC	65594
120	GTAAGATGGTGTGTTGTGAGAGGGCATCAGAGGGCAGACATCTGAAACCACTACACGC	Qy	1259	TTATTTGCCAAATTCAGATTAATTTGAAGAAAGTAGGGAACCACTAGATCACTCAGT	1318
66729	GTAAGATGGTGTGTTGTGAGAGGGCATCAGAGGGCAGATGCACTGMAACCATCTCAC	Db	65593	TTATTTGCCAAATTCAGATTAATTTGAAGAAAGTAGGGAACCACTAGATCACTCAGT	65534
180	AGAAAACTAGTCAATTAATCACTAGGACCAAGCCTTGTCTAATCAATGAACCTAA	Qy	1319	AAAGACCTAAATCCCAATCCCTTATGATTATACAGTGGAAAGTGAGAAATAGATTAAAGGCC	1378
66669	AGAAAGCTAGTCAATTAATCACTAGGACCAAGCCTTGTCTAATCAATGAACCTAA	Db	65533	ATGACCTAAATCCCAATCCCTTATGATTATACAGTGGAAAGTGAGAAATAGATTAAAGGCC	65474
240	GCCATGCCGTGGGCAACCAAGATGGCAGGTCTATGGTGGAGAGATCTGCAGAGATGT	Qy	1379	TAAGTCTGATGACAGAGTACCTAATGAATGAACTATGACAGAGAGTTTCACTGACATTTGACAGG	1438
66509	GCCATGCCGTGGGCAACCAAGATGGCAGGTCTATGGTGGAGAGATCTGCAGAGATGT	Db	65473	TACATCTGATGATAGAGTGCCTGATGAAGTATGGAATGAGGTTTGTGACACTGTACAGG	65414
300	GGTCCACTGGAGAGGGAAT-GCAAAACACTTCAGTATTTCTGCTTTGAGAACCCCATGA	Qy	1439	ACACAGGGATCAGACCACTCCCATGGAAAGAAATGCAAAAAAGCAAAATGGCTGTCTG	1498
66549	GCGCCACCGGAGAGAGATGGCAACCACTTCAGTATTTCTTGCCTTGGAAACCCCATGA	Db	65413	AGACAGGGATCAAGACCACTCCCATGGAAAGAAATGCAAAAAAGCAAAATGGCTGTCTG	65354
359	ACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCCGAGTCAAGTGGT	Qy	1499	GGGAGGCTTCAAAATAGCTGTGAAAGAGAGAGTGAAGCAAGCAAAAGGAGAA	1558
66489	ACAGCATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCCGAGTCAAGTAA	Db	65353	GGGAGGCTTCAAAATAGCTGTGAAAGAGAGAGTGAAGCAAGCAAAAGGAGAA	65294
419	CCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGC	Qy	1559	GATAAAGCAATCTGAATGCAGAGTTCCAAAGAAATAGCAAGAGAGATAGAAAGCCTTCC	65234
66429	CCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGC	Db	1589	-----	1588
479	CAAGCAAAAAGAAATACCCAGCTGTGGATGTGATCTGGTATATGAAGCAAGGTCGATGCT	Qy	65293	GATATAAGCATGTGAATGCAGAGTTCCAAAGAAATAGCAAGAGAGATAGAAAGCCTTCC	65234
66369	CAAGAGAAAACAAATACCCAGCTGTGGATGTGATCTGGTATATGAAGCAAGGTCGATGCT	Db	1589	-----	1588
539	GTAAGAGCAATATTGCATAGGAACCTGGAATGTCAGTCCATGAATCAAGGCAAAATTGG	Qy	65233	TCAGCGATCAATGCAAAAGAAATAGAGGAAACCAACAGAAATGGTAAAGACTAGAGATCTCT	65174
66309	GTAAGAGCAATATTGCATAGGAACCTGGAATGTCAGTCCATGAATCAAGGCAAAATTGG	Db	1589	-----	1588

Db 65173 TCAAGAAATTTAGAGATACCAAGGGGACATTTTCATGCAAAAGATGGGCTTGATTAAGGGCA 65114
 Qy 1589 ----- 1588
 Db 65113 GAAATGGTAGGACCTTAACAGAGAGAGAGATATTAAGAAGAGATGCCAAGAATACACAG 65054
 Qy 1589 ----- 1588
 Db 65053 AAGAACTGTACAAAAAGATCTTCAGACCCAGATAATACAGATGGTGTGATCACTCATC 64994
 Qy 1589 ----- 1588
 Db 64993 TAGAGCAGACATCTGGAATGTGAAGTCAAGTGGCCCTTAGAAGCATCACTACAAACAA 64934
 Qy 1589 ----- 1588
 Db 64933 GCTAGTGGAGGTGATAGAAATTCAGTTGAGCTATTCCAAATCCTGAAAGATGATGCTGTG 64874
 Qy 1589 ----- 1588
 Db 64873 AAGTGTGCACTCAATATGCCAGCAAAATTTGGAATACTCAGCAGTGACCACAGGACTGG 64814
 Qy 1589 ----- 1588
 Db 64813 AAAAGGTCAAGTTTTCATTCOAATCCCAAGAAAGCAATGCCAAAGATGCTCAAACTAC 64754
 Qy 1589 ----- 1588
 Db 64753 CACAAATTCGACTCATCTCAGATGCTAATAAAGTAAATCTCAAAATCTCCAAAGCAGG 64694
 Qy 1589 ----- 1588
 Db 64693 CTTTCAGCAATACGTAACCGTGAATCTCCAGATGTTCAAGCTGGTTTAGAAGAGCAGA 64634
 Qy 1628 GGAACACAGACCAAAATGCAACATCTCTGTATCATGTGAAAGCAAGAGAGATGCCAG 1687
 Db 64633 GGGACCGGAGACCAAAATGCAACATCTCTGTATCATGTGAAAGCAAGAGAGATGCCAG 64574
 Qy 1688 AAAAAATCTATTCTCTGCTTTATGACTATGCAAAAGCCCTTTCAGTGTGGGGTCACAAT 1747
 Db 64573 AAAAAATCTATTCTCTGCTTTATGCGATGCGCAAGCCCTTAAAGTGTGGGTCACAAT 64514
 Qy 1748 AAATGTGGAATTTCTGAAGGATGGGAATACCAAGACCACTGACCTGACTCTTTGAAA 1807
 Db 64513 AAATGTGGAATTTCTGAAGGATGGGAATACCAAGACCACTGACCTGACTCTTTGAAA 64454
 Qy 1808 AATTTGTATCAGGTCAAGGAGCAACAGTTAGAACTGGACATGGAACCAACAGACTGTTTC 1867
 Db 64453 AATCTGTATCAGGTCAAGGAGCAACAGTTAGAACTGGACATGGAACCAACAGACTGTTTC 64394
 Qy 1868 CAAGTAGGAAGAGATGATGCAAGGCTGTATTTGTACCCCGCTTGTTTAACTTTCTAT 1927
 Db 64393 CAATAGGACAAAGGATGATGCAAGGCTGTATTTGTACCCCTGCTATTTAACTTTAT 64334
 Qy 1928 GCAGA-CACATCATGAGAAAGCTGGCTGGAAGAGCACAAGCTGGAATCAAGATGCC 1986
 Db 64333 GCAGTTACATCATGAGAAAGCTGGCTGGAAGAGCACAAGCTGGAATCAAGATGCC 64274
 Qy 1987 GGGAGAAATAGCAATACCTCAGATATGCAAGATGATACCAACCCCTTATGGCAGAAAGTAA 2046
 Db 64273 GGGAGAAATATCAATACCTCAGATATGCAAGATGATACCAACCCCTTATGGCAGAAAGTAA 64214
 Qy 2047 GAGAACTAAAAGCTCTTTGATGAGAGTGAAGAGAGAGTGAAGTGAAGTGGCTTTAAG 2106
 Db 64213 GAGAACTAAAAGCTCTTTGATGAAAGTGAAGAGAGAGTGAAGTGAAGTGGCTTTAAG 64154
 Qy 2107 CTCAATCTCAGAAAAAGAGATCATGGCATCTGGTCCCATCTCATGGGAAATAGAT 2166
 Db 64153 CTGACATTCAGAAAGAGATCATGGCATCCGGTCCCATCTCATGGGAAATAGAT 64094
 Qy 2167 GGGGAAACAGTGGAAACAGTGTGAGACTTTATTTTTTTGGGGGCTCCAAATACATGCGAG 2226

Db 64093 GGGGAAACAGTGGAAAGAGTGGCTGACTTTTA--TTTTGGGGGCTCCAAAATCACTGCGAG 64036
 Qy 2227 ATGGTGAATGACAGCCATGAAATTAAGAGACACTTACTCTTTGGAGAGAAAGTTA--ACCA 2284
 Db 64035 ATGGTGAATGACAGCCATGAAATTAAGAGACACTTACTCTTTGGAGAGAAAGTTATGACCA 63976
 Qy 2285 ACCTAGATAGCATATTTGAAGAGAGACATTTACTCTGCCAACCAAGCCCATCTAGTCA 2344
 Db 63975 ACCTAGATAGCATATTTGAAGAGAGACATTTACTTTTCGGAACAAAGTCTGTCTAGTCA 63916
 Qy 2345 AGGCTATGTTTTCAGTGTGTCATGTATGATGTGAGAGTTGGACTGTGAAGAAAGCTG 2404
 Db 63915 AGGCTATGTTTTCAGTGTGTCATGTATGATGTGAGAGTTGGACTGTGAAGAAAGCTG 63856
 Qy 2405 AGCCTGAAGAAATTTGATCTTTTGAACCTGTGGTGTGTGAGAGACTCTTTGAGAGTCCCTT 2464
 Db 63855 AGCCTGAAGAAATTTGATCTTTTGAACCTGTGGTGTGTGAGAGACTCTTTGAGAGTCCCTT 63796
 Qy 2465 GGACTGCAAGGAGATCCCAACAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCTTTGG 2524
 Db 63795 GGACTGCAAGGAGATCCCAACAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCTTTGG 63736
 Qy 2525 AAGGAATGATCTAAAGCTGAAACTCCAGTACTTTGGCCACTGATCAGAGAGAGTGAAT 2584
 Db 63735 AAGGAATGATCTAAAGCTGAAACTCCAGTACTTTGGCCACTGATCAGAGAGAGTGAAT 63676
 Qy 2585 CACTGGAAGAGCCCTGATGCTGGAGGATTTGGGCGCAGGAGAGAGGAGGAGACAGAGA 2644
 Db 63675 CATTGGAAGAGACTCTGATGCTGGAGGATTTGGGCGCAGGAGAGAGGAGGAGACAGAGA 63616
 Qy 2645 GGATGAGATGGCTGGATGGCATCACTGATCGATGGAGTGAATCTGGGTGAATCTCTGG 2704
 Db 63615 GGATGAGTGGCTGGATGGCATCACTGATCGATGGAGTGAATCTGGGTGAATCTCTGG 63556
 Qy 2705 AGTTGGTATGGACAGGAGGAGGCTGCTCGGGCGATTCATGGGTCAAGAGATTTGA 2764
 Db 63555 AGTTGGTATGGACAGGAGGAGGCTGCTCGGGCGATTCATGGGTCAAGAGATTTGA 63496
 Qy 2765 CACGACTCAGCAACTGAACTGAACTGTA 2797
 Db 63495 CACGACTCAGGAGTGAATGAACTGAACTGAA 63463

RESULT 6
 AC091660
 LOCUS 173068 bp DNA linear MAM 10-OCT-2002
 DEFINITION Bos taurus clone RP42-354B6, complete sequence.
 AC091660
 ACCESSION AC091660.2 GI:15011695
 VERSION HTG.
 KEYWORDS
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos
 1 (bases 1 to 173068)
 REFERENCE
 AUTHORS Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
 Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L.,
 Granite, S., Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E.,
 Lee-Lin, S.-Q., Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B.,
 Masello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J.,
 Pearson, R., Prasad, A., Shevchenko, Y., Snyder, B., Stantripop, S.,
 Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L.,
 Walker, M.A., Wecherby, K.D., Zhang, L.-H. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 173068)
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS


```

TITLE      Direct Submission
JOURNAL    Submitted (25-JUN-2001) NIH Intramural Sequencing Center, 8717
REFERENCE  Grovemont Circle, Gaithersburg, MD 20877, USA
AUTHORS    4 (bases 1 to 173068)
JOURNAL    Green, E.D.
TITLE      Direct Submission
COMMENT    Submitted (10-OCT-2002) NIH Intramural Sequencing Center, 8717
           Grovemont Circle, Gaithersburg, MD 20877, USA
           On Jul 25, 2001 this sequence version replaced gi:14042997.
           ----- Genome Center
           Center: NIH Intramural Sequencing Center
           Center code: NISC
           Web site: http://www.nisc.nih.gov
           Contact: nisc.zoo@nhgri.nih.gov
           ----- Project Information
           Center project name: axa
           Center clone name: 354306

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the Features section.
           Location/Qualifiers
           source
             1..173068
               /organism="Bos taurus"
               /mol_type="genomic DNA"
               /db_xref="taxon:9913"
               /clone="RP42-354B6"
               /clone_lib="RP42"
               /notes="single clone coverage"
           misc_feature
             145040
               /notes="single clone coverage"

Query Match      59.7%; Score 2131; DB 1; Length 173068;
Best Local Similarity 76.6%; Pred. No. 0;
Matches 2628; Conservative 0; Mismatches 165; Indels 636; Gaps 8;

QY 1 GTGAGGAGGGCGGCGGTGAGGAGAGGTACCTCTCGTCCAGGTAAAGGAGCAGTAGC 60
Db 164074 GTGAGGAGGGCGGCGGTGAGGAGAGGTACCTCTCGTCCAGGTAAAGGAGCAGTAGC 164128
QY 61 TGGCTTTGCTGGAGCAGCGCTAAAGAGATACCCACGCCAAGGTAAAGAGAACCCAAAG 120
Db 164129 TGGCTTTGCTGGAGCAGCGCTAAAGAGATACCCACGCCAAGGTAAAGAGAACCCAAAG 164188
QY 121 TAAGATGGTAGGTGTTGTGAGAGGGCATCAGAGGGCAGACATCTGAAACCATACACGA 180
Db 164189 TAAGAGCGTAGGTGTTGCAAGAGGGCATCAGAGGGCAGACATCTGAAACCATACATA 164248
QY 181 GAAAACTAGTCAATCTAATCAGACTAGGACACAGCGCTTGCTAACTCAATGAACTAG 240
Db 164249 GAAAACTAGTCAATCTAATCAGACTAGGACACAGCGCTTGCTAACTCAATGAACTAG 164308
QY 241 CCATGCCCGTGGGCAACCAAGATGGGAGGTTCATGGTGGAGAGATCTGACAGAAATGTG 300
Db 164309 CCATGCCCGTGGGCAACCAAGAGCGGTGGTTCATGGTGGAGAGGTCTGACAGAAATGTG 164368
QY 301 GTCCACTGGAGAGAGGGAAT-GCAAAACCACTTCAGTATCTTGGCTTGAAGAACCCCATGAA 359
Db 164369 GTCCACTGGAGAGAGGGAATGGCAAAACCACTTCAGTATCTTGGCTTGAAGAACCCCATGAA 164428
QY 360 CAGTATCAAGAGGCAAAATGATAGGATCTGAAAGAGGAGTCCCGAGGTCAGTAGTGC 419
Db 164429 CAGTATCAAGAGGCAAAATGATAGGATCTGAAAGAGGAAATTCGCCAGGTCAGTAGTGC 164488

```

```

QY 420 CCCATATGCTACTGGAGATCAGTGGAGAAATAAATCCAGAAAGAAATGAAGAGATGGAGCC 479
Db 164489 CCAGTAGTCTACTGGAGATCAGTGGAGAAATAAATCCAGAAAGAAATGAAGAGATGGAGCC 164548
QY 480 AAGCAAAAGAAATACCAGCTGGATGTGCTGGTGATATAGCAAGTCCGATCCTG 539
Db 164549 AAGCAAAAGAAATACCAGCTGGATGTGCTGGTGATATAGCAAGTCCGATCCTG 164608
QY 540 TAAAGAGCAATATTGCAATAGCAACCTGGAATGTGAGTCCATGAATCAAGCAAAATTTGGA 599
Db 164609 TAAAGAGCAATATTGCAATAGCAACCTGGAATGTGAGTCCATGAATCAAGCAAAATTTGGA 164668
QY 600 AGTGGTCAACAGAGATGGCAAGATGAATGTCAACATTTCTAGGAATCAGCAACTAAA 659
Db 164669 AGTGGTCAACAGAGATGGCAAGATGAATGTCAACATTTCTAGGAATCAGCAACTAAA 164728
QY 660 ATGACTGGAAATGGTGAATTTAACTCAGATGACATTTATATCTTACTTCTGCGGCGAGGA 719
Db 164729 ATGACTGGAAATGGTGAATTTAACTCAGATGACATTTATATCTTACTTCTGCGGCGAGGA 164788
QY 720 ATCCCTCAGAAAGAAATGGAGTAGCATCATGGTCAACAAAGAGTCCGAAATGAGTACT 779
Db 164789 ATCCCTTAGAAGAAATGGAGTAGCATCATGGTCAACAAAGAGTCCGAAATGAGTACT 164848
QY 780 TGGATGCACTCTCAAAAACGACAGATGATCTCTGTTGTTTCCCAAGGCAAAACCAATCAA 839
Db 164849 TGGATGCACTCTCAAAAACGACAGATGATCTCTGTTGTTTCCCAAGGCAAAACCAATCAA 164908
QY 840 TATCACAGTAAATCCAAGTCTATGCCCCCAACAGTAATGCTGAAGAGTGAAGTGAACG 899
Db 164909 TATCACAGTAAATCCAAGTCTATGCCCCCAACAGTAAATGCTGAAGTGAAGTGAACG 164968
QY 900 GTCTATGAGACCTACAGACCTTTTAGAACTAAACCCCAAAAGATGTCCTTCTCAT 959
Db 164969 GTCTATGAGACCTACAGACCTTTTAGAACTAAACCCCAAAAGATGTCCTTCTCAT 165028
QY 960 TATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAAGAACACCTCGAGAGTAACAGGCAAAAT 1019
Db 165029 TATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAAGAACACCTCGAGAGTAACAGGCAAAAT 165088
QY 1020 TGGCTTTGGAATACGGAATGAAGCAGGGCAAGACTAATAGAGTTTTCGCAAGAAATGTC 1079
Db 165089 TGGCTTTGGAATACGGAATGAAGCAGGGCAAGACTAATAGAGTTTTCGCAAGAAATGTC 165148
QY 1080 ACTGTCATAGCAACACCCCTCTTCCAAACACACAGAGAGAGCTCTACACATGACATC 1139
Db 165149 ACTGTCATAGCAACACCCCTCTTCCAAACACACAGAGAGAGCTCTACATATGACATC 165208
QY 1140 ACCAGATGGTCAACACCGAAATCAGATTGATTATATTCTTTGCAAGCCAAAGATGGAGAG 1199
Db 165209 ACCAGATGGTCAACACCGAAATCAGATTGATTATATTCTTTGCAAGCCAAAGATGGAGAG 165268
QY 1200 CTCTATAGCAGTCAAGCAAAACCAAGCAGGAGCTTACTGTGGCTCAGATCATGACTCCT 1259
Db 165269 CTCTATAGCAGTCAAGCAAAACCAAGCAGGAGCTTACTGTGGCTCAGATCATGACTCCT 165328
QY 1260 TATTGCCAAATTCAGACTTAAATTTGAAGAAAGTAGGGAAGAACCCACTAGATCACTCAGGTA 1319
Db 165329 TATTGCCAAATTCAGACTTAAATTTGAAGAAAGTAGGGAAGAACCCACTAGATCACTCAGGTA 165388
QY 1320 AGACTTAAATCCAAATCCCTTTATGATATATACAGTGAAGTGAAGATAGATTTAAGGCGCT 1379
Db 165389 TGACCTAAATCCAAATCCCTTTATGATATATATACAGTGAAGTGAAGATAGATTTAAGGCGCT 165448
QY 1380 AGATCTGATAGCAGAGTACCTTAATGAACACTAGACAGAGTTCATGACTTTGTACAGGA 1439
Db 165449 AGATCTGATAGCAGTGTCTGATGAACACTAGACTGAGGTTTATGATATTTGACAGAA 165508
QY 1440 GACAGGGATCGAGAGCCATCCCATGGAAAGAAATGCAAAAAAGCAAAATGCTCTCTGG 1499
Db 165509 GACAGGGATCAAGACCATCTCTTTGAAAGGAAATGCAAAAAAGCAAAATGCTCTCTGG 165568
QY 1500 GGAGGCGCTTACAAATAGTGTGTAAGAAAGAGAGAGTGAAGCAAGCAAGCAAGAAAGAG 1559

```

Db 165569 GGAGCCCTTACAAATAGCTGTGAAAGAGAGAGAGCAAAAGCAAAGAGAAAGGAAAG 165628
Qy 1560 ATAAAGCATCTGAATGAGAGTTCCAAA----- 1589
Db 165629 ATATAAGCATCTGAATGAGAGTTCCAAAAGAAATAGCAAGAGAGAGATAAGAAAGCCCTTCCCT 165688
Qy 1589 ----- 1588
Db 165689 CAGTGATCAATGCAAGAAATAGAGGAAATAACAGAAATGCGAAGAGCTAGAGATCTCTT 165748
Qy 1589 ----- 1588
Db 165749 CAAGAAATATAGATATACCAAGGAAACATTTTCATGCAAGATGGCTCGATATAAGAGCAG 165808
Qy 1589 ----- 1588
Db 165809 AATGGTATGGA CTTAAGAGAGAGAGAGATATTAATAAGAGGTGGCAGGAATACACAGA 165868
Qy 1589 ----- 1588
Db 165869 AGAATGTACAAAAAGAGCTTCACGACCAGATATACGATGCTGATCACTGACCT 165928
Qy 1589 ----- 1588
Db 165929 AGAGCCAGATCCTGGAATGTGAAGTCAAGTGGGCTTAGAAAGCATCTCTACAAAAA 165988
Qy 1589 ----- 1588
Db 165989 AAAAAAAGAAAGCATCTCTACGAACAAAGCTAGTGGAGGTGATGGAATTCAGTGGAG 166048
Qy 1589 ----- 1588
Db 166049 CTATTTCAAATCCTGAAAGATGATGCTGTGAAAGTGTGCACTCAATATATGCCAGAAAAAT 166108
Qy 1589 ----- 1588
Db 166109 TGGAAACTCAGCAGTGGCCACAGGACTGGAAGGTGAGTTTTCATTTCCATCCAAAA 166168
Qy 1589 ----- 1588
Db 166169 AAAGGCAATGCAAGAAATGCTCAAACTACTGCAAAATTTGACTCATCTCACATGCTAGT 166228
Qy 1589 -----GAACTTCCA 1597
Db 166229 AAGTAATGCTCAAAATTTCCAAAGTAGCTTCAGCAATATGTAACCGTGAACTTCT 166288
Qy 1598 GTTGTTCAGCTGGTTTTAGAAAAGTCAGAGGAACAGAGACCAAAATTTGCCAACATCTCTC 1657
Db 166289 GATGTTCAAGCTGGTTTTAGAAAAGCAGAGGAACAGAGATCAATTTGCCAATATCCAC 166348
Qy 1658 TGATATCATGGAAGAAAGCAAGAGATTCAGAAAAACATCTATTTCTGCTTTTATGACTAT 1717
Db 166349 TGGATCATCAAAAAAGCAAGAGATTCAGAAAAACATCTATTTCTGCTTTTATGACTAT 166408
Qy 1718 GCAAAAGCCCTTGACTGTGGGGTCAATATAACTGTGAAATTTCTGAAAGGATGGGA 1777
Db 166409 GCCAAGCCCTTGACTGTGTGATGATCAATATAAGTGTGAAATTTCTGAAAGAGATGGGA 166468
Qy 1778 ATACCAAGCACTGACTGACTTTGAAAAATTTGTATGCAAGGTGAGGAAGCAACAGTT 1837
Db 166469 ATACCAAGCACCCTGACCTGCTTTGAGAAACCTATATACAGGTGAGGAAGCAACAGTT 166528
Qy 1838 AGAATGGACATGGAACCAAGAGCTGGTTCCAGTAGGAAAGGAGTATGCAAGGCTGT 1897
Db 166529 AGAATGGCCATGGAACCAAGAGCTGGTTCCAAATAGGAAAGGAGTATGCAAGGCTGT 166588
Qy 1898 ATATTGTCAACCGCTTGTTTAACTTCTATGCAAG-ACATCATGAGAAACCGCTGGGTG 1956
Db 166589 ATATTGTCAACCGCTTGTTTAACTTATATGCAAGATACATATGAGAAACACTGGGCTG 166648
Qy 1957 GAAGAGCAACAGCTGGATCAGATTCGGGAGAAATAGCAATACCTCAG-ATATCC 2015

Db 166649 GAAGAGCACAAAGCTGGAATCAAGACTCCCGGATAAATATCAATAAACCCTCAGAATATGC 166708
Qy 2016 AGATGATACCAACCCTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCCTCTTGTATGAAGGT 2075
Db 166709 AGATCAGACGCCCTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCCTCTTGTATGAAGT 166768
Qy 2076 GAAGAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAACGAAGATCATGGC 2135
Db 166769 GAAAGAGAGAGTGAAGAAATTTGGCTTAAAGCTCAACATTCAGAAAAATGAAGATCATGGC 166828
Qy 2136 ATCTCGTCCCATCACTTCATGGGAAATAGATGGGAAACAGTGGAAACAGTGTGAGACTT 2195
Db 166829 ATCTCGTCCCATCACTTCATGGGAAATAGATGGGAAACAGTGGAAACAGTGTGAGACTT 166888
Qy 2196 TATTTTTTGGGGGCTCCAAATCAGTCAGATGGTGAAGTCAAGCCATGAATTAAGA 2255
Db 166889 TATTTTTT--GGGGCTCCAAATCAGTCAGATGGTGAAGTCAAGCCATGAATTAAGA 166946
Qy 2256 CACTTACTCTTTGGAAGAAAAAGTTA--ACCAACCTAGATAGCATATTTGAAAGCAGAGAC 2313
Db 166947 CGCTTCTCTTTGGAAGGAAAGTTATGACCAACCTAGATAGCATATTTCAAAGCAGAGAC 167006
Qy 2314 ATTACCTTGGCAACAAAGCCCATCTAGTCAAGGCTATGTTTTTCCAGTGGTCACTAT 2373
Db 167007 ATTACTTTGGCAACAAAGGCTCATCTAGTCAAGGCTATGTTTTTCTGTGGTCACTAT 167066
Qy 2374 GGATGTGAGAGTGTGAGCTGTGAAGAAAGCTGAGCACTGAAGAAATGATGTTTTTGAAGCTG 2433
Db 167067 GGATGTGAGAGTGTGAGCTGTGAAGAGGCTGAGCAACCAAGAAATGATGTTTTTGAAGCTG 167126
Qy 2434 TGGTGTGGAAGAGACTCTTTGAGAGTCCCTTTGAGACTGCAAGAGAGATCCAAACAGTCCATT 2493
Db 167127 TGGTGTGGAAGAGACTCTTTGAGAGTCCCTTTGAGAGTCTGCAAGGAGATCCAAACAGTCCATT 167186
Qy 2494 CTGAAGAGAGTCAAGCTGGGATTTCTTTGGAAGAAATGATGTTAAAGCTGAAACTCCAG 2553
Db 167187 CTGAAGAGAGTCAAGCTGGGATTTCTTTGGAAGAAATGATGTTAAAGCTGAAACTCCAG 167246
Qy 2554 TACTTTGGCCACCTGATCAAGAGAGCTGACTCACTGGAAGAACCCCTGATGCTGGGAGGG 2613
Db 167247 TAGTTTGGCCACCTCGTGAGAGAGTTGACTCTCTGGAAGAA--CTGATGCTGGGAGGG 167304
Qy 2614 ATTGGGCGCAGAGAGAGAGGAGGAGCAGAGAGATGAGATGGCTGATGCACTGAC 2673
Db 167305 ATTGGGCGCAGAGAGAGAGGAGGAGCAGAGAGATGAGATGGCTGATGCACTGAC 167364
Qy 2674 TCGATGAGCAGTGAAGTCTGGGTGAACTCTCTGAGTGGTGGATGGAACAGGGAGGCTGTCT 2733
Db 167365 TCGATGAGCAGTGAAGTCTGGGTGAACTCTGGGAGCTGGTGGATGGAACAGGGAGGCTGTCT 167424
Qy 2734 CGCGGATTCATGGGTGCAAAAGAGTTGACAGCTGAGCACTGAGCACTGAGCACTGAGCACT 2793
Db 167425 GCTCGGATTCATGGGTGCAAAAGAGTGGCAAGAGTGGCAAGAGTGGCAAGAGTGGCAAGAG 167484
Qy 2794 TGTACTGAA 2802
Db 167485 GGAAGTGA 167493

RESULT 7
AC135215
LOCUS
DEFINITION Bos taurus clone RP42-243A4, WORKING DRAFT SEQUENCE, 3 ordered
pieces
AC135215
VERSION AC135215.2 GI:25699591
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 171153)

Db	36166	TGGATGCAATCTCAAGATGACAGATGATCTCTGTTGGTTTCCAAAGGCAAAACCATTCAG	36225	Qy	1589	-----	1588
Qy	840	TATCACAGTAAATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAAGAAAGCTGAAAGTGAACG	899	Db	37305	TAGTGGAGGTGATGGAATTCACAGTTGAGCTGTTCGAAGTCTCGAAGATAATGCTGGAA	37364
Db	36226	TATCACAGTAAATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAAGAAAGCTGAAAGTGAACG	36285	Qy	1589	-----	1588
Qy	900	GTCCTATGAGAGACCTACAGACCTTTTAGAAGCTTAAACCCCAAAAAGATGCTCCTCTCAT	959	Db	37365	AGTGTGCACTCAATATATCCAGCAAAATTTGGAAAACTCAGCAGTGGCCACAGGACTGGAA	37424
Db	36286	GTTCTATGAAGACCTACAGACCTTTTAGAAGCTTAAACCCCAAAAAGATGCTCCTCTCAT	36345	Qy	1589	-----	1588
Qy	960	TATAGGGAGCTGGAATCCAAAGTAGAAGCAAGAAACACCTGGAGTACAGGCAAAAT	1019	Db	37425	AGGTTCAGTTTTCATTTCCAAATCCCAAGAAAGCAATGCTCCAAACTACCG	37484
Db	36346	TATAGGGAGCTGGAATCCAAAGTAGAAGTCAAGAAACACCTGGAGTACAGGCAAAAT	36405	Qy	1589	-----	1588
Qy	1020	TGGCCTTTGGAAATACGGAATGAAGCAGGGCAAGACTAATAGATTTTGGCCAGAAATGC	1079	Db	37485	CACAATTCACATCTCTCACACGCTAGTAATGTAATGCTCAAAATTCCTCAAGCCAGGCT	37544
Db	36406	TGACCTTTGGAATACAGAATGAAGCAGGGCAAGACTAATAGAGCTTTGGCCAGAAATGC	36465	Qy	1589	-----	1588
Qy	1080	ACTGGTTCATAGCAACACCTCTTCCAAACACACAGAGAGACTCTACACATGGACATC	1139	Db	37545	TCAGCAATATCGGAACCGTGAACCTTCTGTATTTCAAGCTGTTTAGAAAAGTCAGAGG	37604
Db	36466	ACTGGTTCATAGCAACACCTCTTCCAAACACACAGAGAGACTCTACACATGGACATC	36525	Qy	1630	AACCAAGACCAAAATTCGCAACATCTCTGTATCATGGAAAAAGCAAGAGTTTCCAGAA	1689
Qy	1140	ACCAGATGGTCAACACCAAAATCAGATTGATTATATTTCTTTGGCAGCCAAAGATGGAGAAG	1199	Db	37605	AACCAAGATCAAAATTCGCAACATCTCTGTATTTCAAGCTGTTTAGAAAAGTCAGAGG	37664
Db	36526	AGCAGATGGTCAACACCAAAATCAGATTGATTGATTATTTCTTTGGCAGCCAAAGATGGAGAAG	36585	Qy	1690	AAACATCTATTCTGCTTTTATTGACTATGCAAAAGCCCTTTGACTGTGGGGTCAACAATA	1749
Qy	1200	CTCTATACAGTACGCAAAACCAAGACAGAGCTTACTGTGGCTCAGATCATGAATCCT	1259	Db	37665	AAACATCTATTCTGCTTTTATTGACTATGCAAAAGCCCTTTGACTGTGGATCAACAATA	37724
Db	36586	CTCTATACAGTACGCAAAACCAAGACAGAGCTTACTGTGGCTCAGATCATGAATCCT	36645	Qy	1750	ACTGTGAAAAATTTCTGAAAGGGATGGGAATACCAAGACCACCTGACCTGACTCTTTGAAAA	1809
Qy	1260	TATTGCCAAATTCAGACTTAAATTTGAAGAAAGTAGGAAAAACCACTAGATCACTCAGETA	1319	Db	37725	ACCCTGGAATTTCTGAAAGAGATGGGAATACCAAGCCGCTTGATCTGCCTCTTGAGAA	37784
Db	36646	TATTGCC- AATTGAGCTGAAATTTGAAGAAAGCAGGGAACCACTAGACCTTCAGETA	36704	Qy	1810	TTTGTATGCAAGTACGGAAGCAACAGTTAGTAACCTGGACATGGAACAACAGACTGGTTCCA	1869
Qy	1320	AGACCTAAATCCAAATCCCTTATGATTATACAGTGGAAAGTGAAGAAATAGATTTAAGGCT	1379	Db	37785	TTTGTATGCAAGTACGGAAGCAACAGTTAGTAACCTGGACATGGAACAACAGACTGGTTCCA	37844
Db	36705	TGACCTAAATCCAAATCCCTTATGATTATACAGTGGAAAGTGAAGAAATAGATTTAAGGCT	36764	Qy	1870	AGTAGAAAAAGGAGTATGTCAAAGCTGTATATTGTCACCCGCTGTTTAACTTCTATGC	1929
Qy	1380	AGATCTGATAGACAGTACTTAATGAATATGACAGAGGTTTCATGACATGTCACAGGA	1439	Db	37845	AGTAGAAAAAGGAGTACGTCAAAGCTGTATATTGTCACCCGCTGTTTAACTTCTATGC	37904
Db	36765	AGATCTGATAGACAGTACTTAATGAATATGACAGTGGAAAGTGAAGAAATAGATTTAAGGCT	36824	Qy	1930	AGAG-ACATCATGAGAAACGCTGGGCTGGAAAGAGCAACAGCTGGAATCAAGATTGCCGG	1988
Qy	1440	GACAGGATCGAGACCTCCCAATGGAAAAAGAAATGCCAAAAAGCAAAATGGCTGCTCG	1499	Db	37905	AGAGTACATCATGAGAAACCACTGGACTGGAATCAAGAAAAAGCTGGAATCAAGATTGCCGG	37964
Db	36825	GACAGGATCAAGACCATCCCAATGGAAACAGAAATGCAAAAGCAAAATGGCTCTCTGA	36884	Qy	1989	GAGAAATAGCAATAACCTCAGATATGAGATGATACCAACCTTATGCGAGAAAGTGAAGA	2048
Qy	1500	GGAGGCTTCAAAATAGCTGTGAAGAGAGAGTGAAGAGCAAGAAAGAAAGAAAG	1559	Db	37965	GAGAAATAGCAATAACCTCAGATATGAGATGATGACATGACCAACCTTATGCGAGAAATGGA	38024
Db	36885	GGAGGCTTCAAAATAGCTGTGAAGAGAGAGTGAAGAGCAAGAAAGCAAGAAAGAAAG	36944	Qy	2049	GGAACTAAAAAGCCTCTTGATGAAGGTGAAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCT	2108
Qy	1560	ATAAAAGCTCTGAATCGAGATTCCAAA	1588	Db	38025	GGAACTAAAAAGCCTCTTGATGAAGGTGAAAGTGGAGAGTGAAGAAAGTTGGCTTAAAGCT	38084
Db	36945	ATATAACATCTGAATCGAGATTCCAAAGATAGCAAGAGAGATGAAGAAAGCCTTCTT	37004	Qy	2109	CAACATTCAGAAAAAGAGATCATGGCATCTGGTCCCTCATCTTCATGCGGAATAGATGG	2168
Qy	1589	-----	1588	Db	38085	CAACATTCAGAAAAAGTGAAGTATGAGTATGAGATGACCAACCTTATGCGGAATAGATGG	38144
Db	37005	CAGCGATCAATGCAAGAAATAGAGGAAAAACACAGAAATGGGAAAGTAGGATCTCTTCA	37064	Qy	2169	GGAAACAGTGGAAACAGTGTGAGACTTTTATTTTTTGGGGGGCTCCAAAATCACTGCGAGAT	2228
Qy	1589	-----	1588	Db	38145	GGAAACAGTGGAAACAGTGTGAGACTTTTATTTTTT- GTGGGCTCCAAAATCACTACAGAT	38202
Db	37065	AGAAATCAGAGATACCAAGGAACATTTTCATGCAAAAGTAGCTCGATAAGGACAGAA	37124	Qy	2229	GGTGAATCGACCCATGAAATTTAAAGACATTTACTCTTCTTGGAGAAAGTTA- ACCAAC	2286
Qy	1589	-----	1588	Db	38203	GGTGAATCGACCCATGAAATTTAAAGACATTTACTCTTCTTGGAGAAAGTTATGACCAAC	38262
Db	37125	ATGGTATGAATCAACAGAGCAGAGATATTAAGAAAGAGATGGCAAGAAATACACAGAGA	37184	Qy	2287	CTAGATACATATTGAAAAAGCAGAGACATTAACCTTGGCAACAAAGCCCATCTAGTCAAG	2346
Qy	1589	-----	1588	Db	38263	CTAGATACATATTGAAAAAGCAGAGACATTAACCTTGGCAACAAAGCTCCGCTCTAGTCAAG	38322
Db	37185	ACTGTACAAAAAGACCTTCGTGACCCAGATATATGATATGATGTGATCATCTGACCTAGA	37244	Qy	2347	GCTATGGTTTTTTCAGTGGTTCATGTATGGATGTGAGAGTTGGAGTGTGAAGAAAGCTGAG	2406
Qy	1589	-----	1588	Db	38323	GCTATGGTTTTTTCAGTGGTTCATGTAT- GATGTGAGAGTTGGAGTGTGAAGAAAGCTGAG	38381
Db	37245	GCCAGACATCTGGAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTATGAACAAAGC	37304				

Qy	181	GAAACCTAGTCAATCTTAATCACA	TACTAGGACCACAGAGCCTTTGTCTA	ACTCAATGAACCTAAG	240	
Db	80879	GAAACCTAGTCAATCTTAATCACAC	TACTAGGACCAAGAGCCTTTGTCTA	CTCAATGAACCTAAG	80938	
Qy	241	CCATGCCCGTGGGGCAACCCAA	GATGGCAGGTCATGTGTGGAGAGAT	CTGCAGAAATGTG	300	
Db	80939	CCATGCCCGTGGGGCAACCCAAA	ATGGTGCGTCAATGTGTGGAGAGAT	CTGCAGAAATGTG	80998	
Qy	301	GTCCACTGGAGAAGGGAAT-GCAA	CCACTTCAGTATTCTTCCTTGGAA	CCCCCATGAA	359	
Db	80999	GTCCACTGGAGAAGGGAATGGCAA	CCAATTCAGTATTCTTCCTTGGAA	CCCCCATGAA	81058	
Qy	360	CAGTATCAAAAGGCANAATGATA	TAGGNACTGAAAGGAGNACTCCC	CCAGGTCAAGTAGTGC	419	
Db	81059	CAGTATCAAAAGGCANAATGATA	TAGGNACTGAAAGGAGAACTCCC	CCAGGTCAAGTAGTGC	81118	
Qy	420	CCCATATGCTACTGGAGATCAGT	GGAGAAATAACTCCAGAAAGAA	TGAAGATGGAGCC	479	
Db	81119	CCAATATGCTACTGGAGATCAGT	GGAGAAATAACTCCAGAAAGAA	TGAAGATGGAGCC	81178	
Qy	480	AAAGCANAAGAAATACCCAGCT	GTGATGTGATATAAGCAAGGTCC	CCAGTGTG	539	
Db	81179	AAAGCANAAGAAATACCCAGCA	TGTGATGTGATATAAGCAAGGTCC	CCAGTGTG	81238	
Qy	540	TAAAGACAATATTGCATAGAAC	CTCGAATGTCCAGTCCATGAA	TCAAGCAAAATGGA	599	
Db	81239	TAAAGACAATATTGCATAGGAC	CTCGAATGTCCAGTCCATGAA	TCAAGCAAAATGGA	81298	
Qy	600	AGTGGTCAAAACAAGAGATGG	CAAGATGAATTCATPAGAAAT	CAAGCAACTAAA	659	
Db	81299	AGTGGTCAAAACAAGACATGC	AAGAGTGA-----ATTCTAG	AAATCAGCAAACTGAA	81350	
Qy	660	ATGACTTGGAAATGGGTGAAT	TTAACTCAGATGACCAATTATA	CTACTACTCGGGCAGGA	719	
Db	81351	ATGACTTGGAAATGGGTGAAT	TTAACTCAGATGACCAATTATA	CTACTACTCGGGCAGGA	81410	
Qy	720	ATCCCTCAGAAAGAAATGGAG	TAGCCATCATGTGTCAACAA	AGAGTCCGAAATGCACTACT	779	
Db	81411	ATCCCTCAGAAAGAAATGGAG	TAGCCATCATGTGTCAACAA	AGAGTCCGAAATGCACTACT	81470	
Qy	780	TGGATCGACTCTCAAAAACG	ACAGATGATCTGTGTGTTC	CAAGCAACCAATTCAA	839	
Db	81471	TGGATCGAAATCTCAAAAACG	ACAGATGATCTGTGTGTTC	CAAGCAACCAATTCAA	81530	
Qy	840	TATCACAGTAAATCCAAAGT	CTATGCCCCAACCAAGTAAT	CTGGAAGAGCTGAAGTTGAACG	899	
Db	81531	TATCACAGTAAATCCAAAGT	CTATGCCCCAACCAAGTAAC	CTGGAAGAGCTGAAGTTGAACG	81590	
Qy	900	GTCCATGAAGACCTACAGAC	CTTTTAGAATCTACACCC	AAAAAGATGTCTTTCAT	959	
Db	81591	GTTCCTATAAGACCTACAGAA	CTTTTAGAATCTACACCC	AAAAAGATGTCTTTCAT	81650	
Qy	960	TATAGGGACTGGAATGCAAA	AGTAGGAAGCAAGAAACAC	CTGGAGTAAACAGGCAAAAT	1019	
Db	81651	TACAGGGACTGGAATGCAAA	AGTAGGAAGCAAGAAACAC	CTGGAGTAAACAGGCAAAAT	81710	
Qy	1020	TGGCCTTGGAAATACGGAAT	TGAAGCGGCAAGAGCTAAT	TAGAGTTTTGCCAAGAAAAATGC	1079	
Db	81711	TGGCCTTGGAAATACGGAAT	TGAAGCGGCTAAGAGCTAAT	TAGAGTTTTGCCAAGAAAAATGC	81770	
Qy	1080	ACTGGTCTATAGCAAAACAC	CCCTCTTCCCAACACAGAG	AGAGCTCTACATGGACATC	1139	
Db	81771	ACTGGTCTATAGCAAAACAC	CCCTCTTCCAGACACAGAG	AGAGCTCTATATGGACATC	81830	
Qy	1140	ACCAGATGTCTAACCCGAAAT	CAGATTGATTAATCTTTT	GCAGCCAAAGATGGAGAG	1199	
Db	81831	ACCAGATGTCTAACCCGAAAT	CAGATTGATTAATCTTTT	GCAGCCAAAGATGGAGAG	81890	
Qy	1200	CTCTATACAGTCAGCAAAAC	ACAGACCGAGGCTTACTGT	GGCTTCAGATCATGAACTCCT	1259	
Db	81891	CTCTATACAGTCAGCAAAAC	ACAGACCGAGGCTGACTGT	GGCTTCAGACCATGAACTCCT	81950	
Qy	1260	TATTGCCAAATTCAGACTT	TAATTTGAAGAAAGATAGG	AAAAACCACTAGATCACT	CAGGTA	1319

Db	81951	TATTTGCCAAATTCAGACTTTAAATTTGAAGAAAGTAGGGAAAAACACACAGACCATTCAGGTA	82010
Qy	1320	AGACTTAATCCAAATCCCTTTATCATTTATACAGTGGAAAGTGAGAAATAGATTTTAAGGGCT	1379
Db	82011	TGACCTTAATCAAAATCCCTTTATCATTTATACAGTGGAAAGTGAGAAATAGATTTTAAGGGCT	82070
Qy	1380	AGATCTGATAGACAGAGTACCTAATGAACATATGGACAGAGGTTTCATGACATTGTACAGGA	1439
Db	82071	AGATCTGCTAGATAGACTGCCTGATGAACCTATGGAATGAGGTTTCCTGACATTGTACAGGA	82130
Qy	1440	GCACGGGATCGAGACCATCCCATCGAAAAGAAATGCAAAAAGCAAAATGCGTCTCTGG	1499
Db	82131	GCACGGGATCAAGACCATCCCATCGAAAAGAAATGCAAAAAGCAAAATGCGTCTCTGG	82190
Qy	1500	GGAGGCCCTTACAAATAGCTGTGAAAAAGAGAGAAGTGAAAAGCAAAAGGAAAAAGGAAAG	1559
Db	82191	GGAGGCCCTTACAAATAGCTGTGAAAAGAGAAAACGAAAAACAAAGAGAAAAAGGAAAG	82250
Qy	1560	ATAAAAGCATCTGAATGCAGAGTTCCAAA-	1588
Db	82351	ATATAAGCATCTGAATGCAGAGTTCCAAAAGATAGCAAGAGAAATAGAAGAGGCTTCTT	82310
Qy	1589	-----	1588
Db	82311	CAGTGATCAATGCAAGAANAATAGAGGAAAAACAACAGATGGGAAAGACTAGAGATCTCTT	82370
Qy	1589	-----	1588
Db	82371	CAAGAAAAATCAGGGATACCAAAAGGAACATTTTCATGCAAAAGATGGCTCGATAAAGGACAG	82430
Qy	1589	-----	1588
Db	82431	AAATGGTATGGACCTTAAACAGAAAGCAGAGATATTAAAGAGAGATGGCAAGAAATACACAGA	82490
Qy	1589	-----	1588
Db	82491	AGAACTGTATGAAAAGATCTTTCAGCACCAGATAATCATGATGTTGTGTGATCACTGACCT	82550
Qy	1589	-----	1588
Db	82551	AGAGCCAGACATCTTGGAATGTGAAGTCATGTGGGCCCTTTAGAAAGCATCATACGAAACA	82610
Qy	1589	-----	1588
Db	82611	AGCTAGTGGAGGTGATGGAAATCCAGTTGAGCTATTCCAAATCCTGAAAGATGATGCTGT	82670
Qy	1589	-----	1588
Db	82671	GAAAGTGTGCACTCAATATGCCAGCAAAATTTGGAAAACTCAGCAGTGGCCATAGGACTG	82730
Qy	1589	-----	1588
Db	82731	GAAAAGTCAGTTTTCATTCCAAATCCCAAGAAAGGCAATGCCAAAGAAATGCTCAAACTA	82790
Qy	1589	-----	1588
Db	82791	CCACAAATTGCACTCATCTCACATGCTAGTAAAGTAATGCTCAAAATTTTCCAAAGCCAG	82850
Qy	1589	-----GAACTTCCAGTTCTTCAAGCTGTTTGAAGAAAGTCAG	1626
Db	82851	GCTTCAGCAATATGTGAACGTGAACTTCTTGATGTTCAAGCTGGTTTGAAGAAAGGAG	82910
Qy	1627	AGAAACCAAGACCAAAATTCGCAACATCCTCTGTATCATGAAAAAGCAAGAGAGTTCCA	1686
Db	82911	AGAAACCAAGATCAAAATTCGCAACATCCACTGGATCATCGAAAAAGCAAGAGATTTCCA	82970
Qy	1687	GAAAAACATCTATTTCTGCTTTATTGACTATGCAAAAGCCCTTTCGCTGTGGGCTCACAA	1746
Db	82971	GAAAAACATCTATTTCTGCTTTATTGCA-TATGCCAAAGACHTTGAAGTGTGTGGATCACAA	83029
Qy	1747	TAAACTGTGGAAAAATTTCTGAAAGGGATGGGAATACACAGACCACTGACCTGACTTCTTGA	1806

83030	TAACCTGTGAAAAATTTCTGAAAGAGATGGGAATACGAGACCACTGATCTGCTCTTTGAG	83089
1807	AAATTTGTATGCAGGTTCAGGAAGCAACAGTTTAGAACTGGACATGGAAACAACAGACTGGTT	1866
83090	AAATTTGTATGCAGGTTCAGGAAGCAACAGTTTAGAACTGGACATGGAAACAACAGACTGGTT	83149
1867	CCAAGTAGAAAAAGAGTATGTCAAGGCTGTATATTGTCAACCCTGGCTTTTAACTTCTA	1926
83150	CCTAAATAGAAAAGAGTACGTCAAGGCTGTATCTGTAACTCTGTTTATTTCCTCTATA	83209
1927	TGCAGAG-ACATCATGAGAAACGCTTGGCTTGGAAAGAAAGCAACAAGCTGGAATCAAGATTGC	1985
83210	TTTCAAGTACATCATGAGAAACGCTGGACTGGAGAGAGCAACAAGCTGGAATCAAGATTGC	83269
1986	CGGGAGAAATAGCAATAAACCCTCAGATATGCGATGATGATACCAACCTTATGGCAGAAAGTGA	2045
83270	CGGGAGAAATATCAATAAACCCTCAGATATGCGATGATGATGATACCAACCTTATGGCAGAAAGTGA	83329
2046	AGAGGAACATAAAGAGCCTCTTCATCAAGGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAA	2105
83330	AGAGGAACTATAAAGAGCCTCTTCATCAAGGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAA	83389
2106	GCTCAACATTTCAGAAAAAGAGATCATGGCATCTGGTCCCATCTTCATGGGAAATAGA	2165
83390	GCCCAACATTTCAGAAAAAGAGATCATGGCATCTGGTCCCATCTTCATGGGAAATAGA	83449
2166	TGGGGAAAAGTGGAAAACAGTCTCAGACTTTATTTTTTTGGGGGGCTCCAAATCACATGCA	2225
83450	TGGGGAAAAGTGGAAAACAGTCTCAGACTTTATTTTTTTCTGGGCTCCAAATCATGCA	83507
2226	GATGGTGACTGCAGCCATCAAAATTTAAAAGACACTTACTCTTTGGAAGAAAAGTTA--ACC	2283
83508	GATGGTGACTGCAGCCATCAAAATTTAAAAGATGCTTTGCTCTTGGAAAGAAAGTTATGACC	83567
2284	AACTAGATAGCATTTGAAAAGCAGAGACATTACTTTGCCAACAAAGCCCATCTAGTC	2343
83568	AACTAGATAGCATTCTAAAAGCAGAGACATTACTTTGCCAACAAAGGTCGTCTAGTC	83627
2344	AAGGCTATGGTTTTTCCAGTGGTCAATGTATGGATGTGAGAGTTGACCTGTCAAGAAAGCT	2403
83628	AAGGCTATGGTTTTTCCAGTGGTCAATGTATGGATGTGAGAGTTGGCTGTCAAGAGGCT	83687
2404	GAGCATTGAAGAAATGATGCTTTTTGAACTGTGGTGTGGAGAGACTCTTGAGAGTCCCT	2463
83688	GAGCG-CCGAAGAAATGATGCTTTTTGAACTGTGGTGTGGAGAGACTCTTGAGAGTCCCT	83747
2464	TGGAC-TGCAAGAGATCCAAACAGTCCATCTCTGAAGGAGATCAAGCTGGGATTTCTTT	2522
83748	TGACATTGCAAGAGATCCAAACAGTCCATCTCTGAAGGAGATCAAGCTGGGATTTCTTT	83807
2523	GGAGGAATGATGCTTAAAGCTGAACTCCAGTACTTTGGCCACCTGTATCAAGAGGCTGA	2582
83808	GGAGGAATGATGCTTAAAGCTGAACTCCAGTACTTTGGCCACCTGTATCAAGAGGCTGA	83867
2583	CTCACTGAAAAAGACCTGTATGCTGGAGGGATTTGGGCGAGGAGAGAGGGGACGACA	2642
83868	C-CCTTGGAAAAAGACTCTGATCTCTGGAGGGATTTGGGGCGAGAGAGAGGGGACGACA	83926
2643	GAGGATCAGATGGCTGATGGGATCACTGACTCGATGGACGTTGAGTCTGGGTGAACCTCCT	2702
83927	GAGGATCAGATGGCTGATGGGATCACTGACTCGATGGACGTTGAGTCTGGGTGAACCTCCT	83986
2703	GGAGTTGGTGAATGGAACAGGAGGCGCTGTCTCTGGGGGATTCATGGGGTCAAAAGAGTTG	2762
83987	GTAGTTGGTGAATGGAACAGGAGGCGCTGTCTCTGGGGGATTCATGGGGTCAAAAGAGTTG	84046
2763	GACACGACTGAGCAACTGAACTGAACTGAACTGTACTGTAGTAAACCTTAGTAGTTTATATAC	2822
84047	GACACGACTGAGCAACTGAACTGAACTGTACTGTAGTAAACCTTAGTAGTTTATATAC	84106
2823	TCGAAAAATAGTAATTTTATATGTATCAAAATTTATTTCAATATGTTTGGTTAAGATAATA	2882
84107	GACTTTTCAATATATATATTTGATTTTATTTTCAAACTCTATGGGATAGATTGTT	84166

Qy	2883	AGATTTTCAAATGATTTTTATCTTTGATTTTTCTCTACTTAATTAAATTTTGGGATTTTA	2942
Db	84167	AATACTCTTTTACAGTTGAGGAAATTTGATGCCAAAAGCTGAATGACTTTTCCAGGGTCT	84226
Qy	2943	ACATTTCTTCAATGACTTGTATTTCTTAATATTTACTTATTTACTTTTACTTTTA	2996
Db	84227	AATTACTAATAAATACTCGCAGAACTAAGATATATATAAGGTACTCTGGATTTCA	84280

RESULT 9	AC141853	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
----------	----------	-------	------------	-----------	---------	----------	--------	----------

AC141853 192431 bp DNA linear HTG 12-APR-2003
 Bos taurus clone RP42-374H3, WORKING DRAFT SEQUENCE, 3 ordered
 pieces.
 AC141853
 AC141853.2 GI:29824448
 HTG: HTGS PHASE2: HTGS DRAFT.

SOURCE
BOS taurus (cow)
BOS taurus
BOS taurus
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
Bovidae; Bovinae; Bos.
REFERENCE
1 (bases 1 to 12431)
AUTHORS
Antonellis, A., Avele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

REFERENCE AUTHORS

1 (bases 1 to 192431)
 Evidence: 1909
 Antonellis, A., Ayete, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Carriage, K.,
 Chu, G., Coleman, B., Coleman, H., Engle, J., Granate, S., Guan, X.,
 Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,
 Hurle, B., Icol, J.R., Karlins, E., Kwong, P., Laric, P., Lee-lin, S.-Q.,
 Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Mastello, C.,
 Maskati, B., McDowell, J., Pagniragan, C., Pearson, R., Portnoy, M.E.,
 Prasad, A., Reddik-Dugue, N., Schandler, K., Schueler, M.G., Shah, K.,
 Sison, C., Svanthorp, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,
 Yeast, T., Zwanberg, K.D., Wiggins, L., Young, A. and Great E.D.

TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
1				2				3	
4				5				6	
7				8				9	
10				11				12	
13				14				15	
16				17				18	
19				20				21	
22				23				24	
25				26				27	
28				29				30	
31				32				33	
34				35				36	
37				38				39	
40				41				42	
43				44				45	
46				47				48	
49				50				51	
52				53				54	
55				56				57	
58				59				60	
61				62				63	
64				65				66	
67				68				69	
70				71				72	
73				74				75	
76				77				78	
79				80				81	
82				83				84	
85				86				87	
88				89				90	
91				92				93	
94				95				96	
97				98				99	
100				101				102	
103				104				105	
106				107				108	
109				110				111	
112				113				114	
115				116				117	
118				119				120	
121				122				123	
124				125				126	
127				128				129	
130				131				132	
133				134				135	
136				137				138	
139				140				141	
142				143				144	
145				146				147	
148				149				150	
151				152				153	
154				155				156	
157				158				159	
160				161				162	
163				164				165	
166				167				168	
169				170				171	
172				173				174	
175				176				177	
178				179	</				

NIH Intramural Sequencing Center, 8717
Grosvont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 192431)
Direct Submission
Submitted (19-MAR-2003) NIH Intramural Sequencing Center, 8717
Grosvont Circle, Gaithersburg, MD 20877, USA
2 (bases 1 to 192431)
Unpublished
NIH Comparative Sequencing Initiative
Green, E.D.
2 (bases 1 to 192431)

COMMENT

Government Circle, Gaithersburg, MD 20877, USA
 On Apr 12, 2003 this sequence version replaced gi:29124103.

 Center: Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@hghri.nih.gov

 Project Information

 Center project name: djp
 Center clone name: 374H03

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 192108 bases at least Q40

Consensus quality: 192184 bases at least Q30
 Consensus quality: 192227 bases at least Q20

Insert size: 188000; agarose-fp
 Insert size: 192231; sum-of-contigs

Quality coverage: 11-22x in Q20 bases; agarose-fp
 Quality coverage: 10.97x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 24496: contig of 24496 bp in length
 * 24497 24596: gap of unknown length
 * 24597 47566: contig of 22970 bp in length
 * 47567 47666: gap of unknown length
 * 47667 192431: contig of 144765 bp in length.

FEATURES

source	1. 192431
	/organism="Bos taurus"
	/mol_type="genomic DNA"
	/db_xref="taxon:9913"
	/clone="RP42-374H3"
	/clone_lib="RP42"
misc_feature	1. 24496
	/note="assembly_fragment"
	clone end:SP6
	vector side:left"
misc_feature	24597..47566
	/note="assembly_fragment"
misc_feature	47667..192431
	/note="assembly_fragment"
	clone end:T7
	vector side:right"
misc_feature	85439..192431
	/note="clone overlaps with GenBank Accession Number AC141852 clone RP42-269D20 (center project name djf)"
Query Match	59.4%; Score 2123.49; DB 1; Length 192431;
Best Local Similarity	75.1%; Pred. No. 0;
Matches 2699; Conservative	0; Mismatches 280; Indels 615; Gaps 10;
QY	1 GTCCAGGAGGCGCAGTGGAGGAGGTACCTACCTCGTCCNAGGTAAGGACGATGC 60
Db	29026 GTCCAGGAGGCGGCTGGTGGAGGA-----TACCCCTCTCCCAAGGTAAGGACATGGC 29080
QY	61 TGGCGTTTGTGGAGAGCGCGTAAAGATACCCACGCCCAAGGTAAGAGAAACCCAG 120
Db	29081 TGGCTTTGTGGAGCGCTGTGAAGATACCCACGCCCAAGGTAAGAGAAACCCAG 29140
QY	121 TAAGATGGTGGTGGTGGAGGGGATCAGAGGGGAGACATCTGAACCCATACACGCA 180
Db	29141 TAAGACGGTAGGTGGTGGAGAGGGGATCAGAGGGGCAACACACTGAACACCATCTATA 29200
QY	181 GAAAACTAGTCAATCTAATCACACTAGGACCAAGCCTTGTCTAATCTCAATCAACTAAG 240
Db	29201 GAAAACTAGTCAATCTAATCACACTAGGACCAAGCCTTGTCTAATCTCAATCAACTAAG 29260
QY	241 CCATGCCGTGGGGCAACCCAGATCGGAGGTCTATGGTGGAGAGATCTGACAGATGTG 300
Db	29261 CCATGCCGTGGGGCAACCCAGATCGGAGGTCTATGGTGGAGAGATCTGACAGATGTG 29320
QY	301 GTCCACTGGAGAGGGAAT-GCAAAACCACTTCAGTATCTTCCCTTGGAGCCCATGAA 359
Db	29321 GTCCACTGGAGAGGGAATGGCAACCACTTCAGTATCTTCCCTTGGAGCCCATGAA 29380
QY	360 CAGTATGAAAAGGCAAAATGATAGGATCTGAAAGAGGAATCCCGAGGTCAAGTGGTGC 419
Db	29381 CAGTATGAAAAGGCAAAATGATAGGATCTGAAAGAGGAATCCCGAGGTCAAGTGGTGC 29440

QY	420 CCCATATGCTACTGAGATCAGTGGAGAAATAACTCCAGAAAGATGAAGATGGAGCC 479
Db	29441 CCAATATGCTACTGAGATCAGTGGAGAAATAACTCCAGAAAGATGAAGATGGAGCC 29500
QY	480 AAGCAAAAAGAAATACCCAGCAGTGGATGTGACTGGTGCATATAAGCAAGGTCCGATGCTG 539
Db	29501 AAGCAAAAAGAAATACCCAGCAGTGGATGTGACTGGTGCATATAAGCAAGGTCCGATGCTG 29560
QY	540 TAAAGAGCAATATTCATAGGAACCTGGAATCTCAGGTCCATGATCAAGCAAAATGGA 599
Db	29561 TAAAGAGCAATATTCATAGGAACCTGGAATCTCAGGTCCATGATCAAGCAAAATGGA 29620
QY	600 AGTGGTCAAAAGAGATGCGAAGATGGAATGTCAACATTTCTAGGAATCAGCAACTAAA 659
Db	29621 AGTGGTCAAAAGAGATGCGAAGATGGAATGTCAACATTTCTAGGAATCAGCAACTAAA 29672
QY	660 ATGGACTGGAATGGTGAATTTAACTCAGATGACCATTTATATCTACTCTCGGGCAGGA 719
Db	29673 ATGGACTGGAATGGTGAATTTAACTCAGATGACCATTTATATCTACTCTCGGGCAGGA 29732
QY	720 ATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCGAAATGCACTACT 779
Db	29733 ATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCTGAATGCACTACT 29792
QY	780 TGGATGCACTCTCAAAAACGACAGATGATCTCTGTTGTTTCCAAAGCAAAACCATCAA 839
Db	29793 TGGATGCACTCTCAAAAACGACAGATGATCTCTGTTGTTTCCAAAGCAAAACCATCAA 29852
QY	840 TATCACAGTAATCAAGTCTATGCCCAACAGTAATCTGAAAGAGTGAAGTTGAACG 899
Db	29853 TATCACAGTAATCAAGTCTATGCCCAACAGTAATCTGAAAGAGTGAAGTTGAACG 29912
QY	900 GTCTTATGAAGACCTTACAAGACCTTTTAGAACTAACCAACCAAAAGATGTCTCTTCAT 959
Db	29913 GTCTTATGAAGACCTTACAAGACCTTTTAGAACTAACCAACCAAAAGATGTCTCTTCAT 29972
QY	960 TATAGGAGTCTGGAATCGAAGTAGGAAGCAAGCAAGCAACCTGGAGTACAGCAATTT 1019
Db	29973 TATAGGAGTCTGGAATCGAAGTAGGAAGCAAGCAACCTGGAGTACAGCAATTT 30032
QY	1020 TGGCTTTGGAAATACGGAATGAAGCAGGCGCAAGACTAATAGAGTTTTCACCAAGAAATGC 1079
Db	30033 TGGCTTTGGAAATACGGAATGAAGCAGGCGCAAGACTAATAGAGTTTTCACCAAGAAATGC 30092
QY	1080 ACTGGTCAATAGCAACCCCTCTTCCAAACACCAAGAGAGACTCTACACATGGAATC 1139
Db	30093 ACTGGTCAATAGCAACCCCTCTTCCAAACACCAAGAGAGACTCTACACATGGAATC 30152
QY	1140 ACCAGATGGTCAACACCGAAATCAGATTGATTATTTCTTTGACGCCAAAGATGGAGAAG 1199
Db	30153 ACCAGATGGTCAACACCGAAATCAGATTGATTATTTCTTTGACGCCAAAGATGGAGAAG 30212
QY	1200 CTCTATACGTACAGCAAAAAACAAGCCAGGAGTCTACTGGCTCAGATCATGAATCTCT 1259
Db	30213 CTCTATACGTACAGCAAAAAACAAGCCAGGAGTCTACTGGCTCAGATCATGAATCTCT 30272
QY	1260 TATTGCCAAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTA 1319
Db	30273 TATTGCCAAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTA 30332
QY	1320 AGACTTAATCCATCCCTTATGATTATACAGTGGAGAGTGAAGAAATAGATTTAAGGCGCT 1379
Db	30333 AGACTTAATCCATCCCTTATGATTATACAGTGGAGAGTGAAGAAATAGATTTAAGGCGCT 30392
QY	1380 AGATCTGATAGACAGAGTACCTAATGAACTATGGAAGAGAGTTCATGACATTTGACAGA 1439
Db	30393 AGATCTGATAGACAGTACCTAATGAACTATGGAAGAGAGTTCATGACATTTGACAGA 30452
QY	1440 GACAGGATCGAGACCATCCCGATGAAAGAAATGCAAAAAAGCAAAATGGCTGTCTGG 1499
Db	30453 GACAGGATCGAGACCATCCCGATGAAAGAAATGCAAAAAAGCAAAATGGCTGTCTGG 30512

QY	1500	GGAGGCTTACAAATAGCTGTGAAAGAGAGAGAGTGAAGGCAAGGAAAGGAAAG	1559
Db	30513	GGAGGCTTACAAATAGCTGTGAAAGAGAGAGAGTGAAGGCAAGGAAAGGAAAG	30572
QY	1560	ATAAAGCATCTGAATGCGAGTTCAAA	1588
Db	30573	ATATAAGCATCTGAATGCGAGTTCAAAAGAAATAGCAAGAGAAATPAAGAAAGGCTTCTT	30632
QY	1589	-----	1588
Db	30633	CAGTGATCAATGCAAGAAATAGAGAAACACAGAAATGGGAAAGACTAGAGATCTCTT	30692
QY	1589	-----	1588
Db	30693	CAAGAAAATCAGGATACCAAGAGCAATTTTCATGCAAGATGGGCTCGATTAAGGACAG	30752
QY	1589	-----	1588
Db	30753	AAATGGTATGGACCTAACAGAGCAGAGATATTAAGAGAGATGGCAAGAAATACACAGA	30812
QY	1589	-----	1588
Db	30813	AGAACTGATGAAAGATCTTCACGACCCAGATAATCATGATGGTGTGATCACTGACCT	30872
QY	1589	-----	1588
Db	30873	AGAGCCAGACATCCTGGAATGTGAAGTTCATGTGGGCTTTAGAAAGCATCACTACGAAACAA	30932
QY	1589	-----	1588
Db	30933	AGCTAGTGGAGTGATGGAATTCAGTTGAGCTATTCCAAATCCTGAAAGATGATGCTGT	30992
QY	1589	-----	1588
Db	30993	GAAAGTCTGCATCAATATATCCAGCAAAATTTGGAAAATCTCAGCAGTGGCCATGAGACTG	31052
QY	1589	-----	1588
Db	31053	GAAAGTCTAGTTTTCATTCATCCAAAGAAAGGCAATGCCAAAGATGCTCAAACTA	31112
QY	1589	-----	1588
Db	31113	CCACAAATTCGACTCTCACATCTAGTAAAGTAATGCTCAAAATTTTCCAAAGCCAG	31172
QY	1589	-----	1626
Db	31173	GCCTCAGCAATATGTGAATCTGTGAATCTCCTGATGTTCAAGCTGGTTTGTAGAAAGGCG	31232
QY	1627	AGGAACGAGACCAAAATTTGCCAACATCTCTGTATCATGGAAAAAGCAAGAGAGTTCCA	1686
Db	31233	AGGAACGAGATCAATTTGCCAACATCTCTGTATCATGGAAAAAGCAAGAGATTTCCA	31292
QY	1687	GAAAAACATCTATTCTCTTTATTTGACTATGCAAAAGCCCTTTGACTGTGGGGTCAAA	1746
Db	31293	GAAAAACATCTATTCTCTTTATTTGA-TATGCAAAAGACTTTGAGTGTGGATCAAA	31351
QY	1747	TAAACTGTGGAAATCTGAAGGATGGGAATACCAAGACACCTGACCTGACTCTTGAA	1806
Db	31352	TAAACTGTGGAAATCTGAAGGATGGGAATACCAAGACACCTGATCTGCCCTCTTGAG	31411
QY	1807	AAATTTGATGACAGTCAGGAAGCAACAGTTTAGAATCTGGACATGGAAACACAGACTGGTT	1866
Db	31412	AAATTTGATGACAGTCAGGAAGCAACAGTTTAGAATCTGGACATGGAAACATCAGACTGGTT	31471
QY	1867	CCAAAGTAAAGAGATGTCAAGGCTGTATATGTCAACCCGCTTGTATTAACTCTA	1926
Db	31472	CCTAATAGAAAGAGTACGTCAAGGCTGTATATGTCAACCCGCTTGTATTTCCTTATA	31531
QY	1927	TGCAGAG-ACATCATGAGAAACGCTGGGCTGGAAAGCAACAAAGCTGGAATCAAGATTGC	1985
Db	31532	TTCAGGTACATCATGAGAAAGCTGGGCTGGAAAGCAACAAAGCTGGAATCAAGATTGC	31591
QY	1986	CGGGAGAAATAGCAATAACCTCAGATATGCAATATGCAATACCACTTATGGCAGAAAGTGA	2045

RESULT 10

AC108887

LOCUS

DEFINITION

AC108887 173207 bp DNA linear HTG 15-AUG-2002
Bos taurus clone Rp42-14F6, WORKING DRAFT SEQUENCE, 4 ordered
pieces.

Db	31592	CGGAGAAATCAATAACCTCAGATATGCACCACTTAATGCGAGAAAGTGA	31651
QY	2046	AGAGAACTAAAGACCTCTTTGATGAAGTGAAGAGAGAGAGTGAAGAAAGTTGCTTAAA	2105
Db	31652	AGAGAACTAAAGACCTCTTTGATGAAGTGAAGAGAGAGAGTGAAGAAAGTTGCTTAAA	31711
QY	2106	GCTCAACATTGAGAAAAAGAGATCATGCATCTGGTCCCATCACTTCATCGGAAATAGA	2165
Db	31712	GCCCAACATTGAGAAAAAGAGATCATGCATCTGGTCCCATCACTTCATCGGAAATAGA	31771
QY	2166	TGGGAAACAGTGGAAACAGATGTCAGACTTTATTTTGGGGGGCTCCAAAATCACTGCA	2225
Db	31772	TGGGAAACAGTGGAAACAGATGTCAGACTTTATTTTCTGGGCTCCAAAATCACTGCA	31829
QY	2226	GATGTGATCTGAGCCATGAAATTAAGACACTTACTCTCTTGGAAAGAAAGTTA--ACC	2283
Db	31830	GATGTGATCTGAGCCATGAAATTAAGACACTTACTCTCTTGGAAAGAAAGTTATGACC	31889
QY	2284	AACCTAGATAGCATATTGAAAGAGAGACATTAACCTTGCCCAACAAAGCCCATCTAGTC	2343
Db	31890	AACCTAGATAGCATATTGAAAGAGAGACATTAACCTTGCCCAACAAAGGTCTGTCTAGTC	31949
QY	2344	AAGGCTATGTTTTCAGTGGTCAATGATGATGAGAGTTGAGCTGTGAAGAAAGCT	2403
Db	31950	AAGGCTATGTTTTCAGTGGTCAATGATGATGAGAGTTGAGCTGTGAAGAAAGCT	32009
QY	2404	GAGCACTGAGAAATTTGATGCTTTTGAACCTGTGGTGTGGAGAAAGCTCTTGAGAGTCCCT	2463
Db	32010	GAGCCCGGAAATTTGATGCTTTTGAACCTGTGGTGTGGAGAAAGCTCTTGAGAGTCCCT	32069
QY	2464	TGGAC-TGCAAGAGAGATCCAAACAGTCCATCTGAAGGAGATCAGCCCTGGGATTTCTTT	2522
Db	32070	TGGACTTGAAGAGATCCAAACAGTCCATCTGAAGGAGATCAGCCCTGGGATTTCTTT	32129
QY	2523	GGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGAAAGCTGA	2582
Db	32130	GGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGAAAGCTGA	32189
QY	2583	CTCACTGGAAGACCTGATCTGGGAGGATTTGGGCGAGGAGAGAGGAGGAGGAGACA	2642
Db	32190	C-CTTTGGAAAAAGACTCTGATCTGGGAGGATTTGGGCGAGGAGAGAGGAGGAGGAGACA	32248
QY	2643	GAGGATGAGATGGCTGGATGGCATCACTGATCGATGAGAGCTGAGTCTGGGTGAAGCTCT	2702
Db	32249	GAGGATGAGATGGCTGGATGGCATCACTGATCGATGAGAGCTGAGTCTGGGTGAAGCTCT	32308
QY	2703	GAGTTGGTGTGAGCAGGAGGAGGCTGCTCGCGGATTCATGGGGTCAACAAAGTTG	2762
Db	32309	GAGTTGGTGTGAGCAGGAGGAGGCTGCTCGCGGATTCATGGGGTCAACAAAGTTG	32368
QY	2763	GACAGACTGAGCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT	2822
Db	32369	GACAGACTGAGCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT	32428
QY	2823	TCAGAAAAATGATTAATTCATATGATTAATTCATTAATTCATTAATTCATTAATTCATTAAT	2882
Db	32429	GACTTTCATATATATATGATTAATTCATTAATTCATTAATTCATTAATTCATTAATTCATTAAT	32488
QY	2883	AGATTTTCAATGATTTTATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2942
Db	32489	AATACTCTTTTCAAGTTGAGGAAATTTGATGAGGAAAGCTGAACTGAACTTTCCAGGCTCT	32548
QY	2943	ACTATTCTTCAATGACTTGTATTTCTAATATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	2996
Db	32549	AATTACTAATAAATACTGCAAGAACTAAGATATATATAGGTATCTGGAATCA	32602


```
misc_feature      /clone lib="RP42"
1.1801
/notes="assembly_fragment
clone_end:T7
vector_side:left"
misc_feature      1902..41677
/notes="assembly_fragment"
41778..81698
/notes="assembly_fragment"
81799..95185
/notes="assembly_fragment"
95286..106702
/notes="assembly_fragment"
106803..152491
/notes="assembly_fragment"
129729..213549
/notes="clone overlaps with GenBank Accession Number
AC140962 clone RP42-329016 (center project name dju)"
152592..213549
/notes="assembly_fragment
clone_end:SP6
vector_side:right"

Query Match      59.0%; Score 2107.7; DB 1; Length 213549;
Best Local Similarity 86.1%; Pred. No. 0;
Matches 2504; Conservative 0; Mismatches 243; Indels 161; Gaps 17;

Qy 22 GAGAGTACCTACCTCGTCCAGGTAAGGAGCAGTAGCTGCGTTTGCGGAGCGCG 81
Db 24602 GAGAGGATACCCCGCTCCAGGTAAGGAGCAGCGCGCTTTGCTGGAGCGCCC 24661

Qy 82 -TAAAGAGATACCCCGCCAGGTAAGGAGAACCCCAAGTAAGATGTTGGTGA 140
Db 24662 ATGAAGAGACACCATGTCCAGATGAAGAAACCCCACTAGACGTTAGTGTGAGA 24721

Qy 141 GAGGCGATCAGAGCGCAGACATCTGAAACCATACACGCGAGAAACTAGTCAATCTAATC 200
Db 24722 CAGGCGATCAGAGCGCAG- - -ACTGAAACCAACATCACAGAAACTTAGCCATCTAATC 24777

Qy 201 ACATAGGACACAGCCTTGTCTAATCAATGAATCAAGCCATGCC -CGTGGGCAACC 259
Db 24778 ACA -TGGACATAGCCTTATCTAATCAATGAACCAAGCATGTCATATAGCGCCAC 24835

Qy 260 CAAGATGGCAGGTCATGGTGGAGAGATCTGACAGAAATGTGTCCTCACTGGAGAGGGAAT 319
Db 24836 CAAGACGTCAGGTCACGGTGGAGAAATCTGACAAATTTGGTCCACTGGAGAGGGAAT 24895

Qy 320 -GCAAAACCACTTACATTTCTTGGCTTGGAGAACCCCATGAACAGTATGAAAGGCAAAAT 378
Db 24896 GGCAAACCACTTCAATATTTCTTGGCTTGGAGAAATCCCATGAATAGTATGAAAGGCAAAA 24955

Qy 379 GATAGGATCTGAAGAGGAATCCCGAGGTCAGTGTGCCCCCATATGCTACTGGAGAT 438
Db 24956 GATAGGACACTGAAGATGAATCTCCCAAGTGGACAGTGTGCCCATATGCTACTGGAGAT 25015

Qy 439 CAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCCAAAGCAAAA- - - 488
Db 25016 CAGTGGAGAAATAACTCCAGACAGAAATGAAGAGCAGAGCCAAAGCAAAAACAA 25075

Qy 489 - - -AGAAATACCCAGTGTGGATGTGACTGGTGATATGAAGCAGGTCCGATGCTGAAGA 545
Db 25076 AACAAAAACCCAGTGTGGATGTGACTGGTGATATGAAGCAGGTCCCAATGCTGTGAAGA 25135

Qy 546 GCAATATTGATAGGAACCTGGAATGTGAGTCCATGAATCAAGGCAAAATTTGGAAGTGGT 605
Db 25136 GCAATATTGATAGGAACCTGGAATGTGAGTCCATGAATCAAGGAAATTTGGAAGTGG- 25194

Qy 606 CAACAGAGATGGCAAGTAGTATGTCAAATTTCTAGGAATCAGCGAACTAAATGGAC 665
Db 25195 CAATAGAGATGGCAAGTAGTAAATCAACATTTCTAGGAATCAG -TAACTAAGATGGAC 25253

Qy 666 TGGATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTCGCGGAGGAATCCCT 725
```

```
Db 25254 TGCAATGGGTGAACCTTAATCAGATGACCATTATATCTACTATATAGTGGCAGAAATCCCT 25313
Qy 726 CAGAAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCGAAATGCGACTTCTGGATG 785
Db 25314 TAGAAGAAATTAATAGCCATCATAGTCAACAAAGAAATCCAAATGCGACTTCTGGATG 25373
Qy 786 CAGTCTCAAAACGACAGAAATGATCTCTGTTGTTTCCAAAGGCAAAACCAATCAATATCAC 845
Db 25374 TAATCCC -AAATGACAGATGATCCTGCTGCTTTCCAAAGGCAAAACCAATCAATATCAC 25432
Qy 846 AGTAATCCAAGTCTATGCCCAACCAAGTAATGCTGAAGAACTGAAGTTGAACGGTCTCTA 905
Db 25433 AGTAATCCAAGCTATGCCCAACCAAGTAACGCTGAAGAACTGAAGTTGAATGCTCTG 25492
Qy 906 TGAAGACCTACAGACCTTTTGAAGACTAACCCCAAAAGAGTGTCTTCTCATATAGG 965
Db 25493 TGAAGACCTACAGACCTCTTTAGAACTAACCCCTAAAAAAGATGTCTTTTCAITATAGG 25552
Qy 966 GGACTTGAATCCAAAGTAGGAAGCAAAAGCAACCTCGGAGTTAACAGGCAAAATTTGGCCT 1025
Db 25553 GAACCTGGAAATGCANAAAGTGGGAACTCAAGAAACCTCGGAGTTAACAGGCAAAATTTGGCCT 25612
Qy 1026 TGAATAGCGAATGAAGCAGGCAAGAACTAATAGAGTTTTCGCAAGAAATGCACTGGT 1085
Db 25613 TGAATAGCGAATGAAGCAGGCAAGAACTAATAGAGTTTTCGCAAGAAATGCACTGGT 25672
Qy 1086 CATAGCAACACCCCTTCCCAACACACAGAGAGAACTCTACACATGCAATCACCAGA 1145
Db 25673 CATAGCAACACCCCTTCTCCAAACACAC - - -AAGGCTCTACATGAACTCACCAGA 25727
Qy 1146 TGGTCAACACCGAAATCAGATTGATTATTTCTTTCAGGCAAGAAATGGAAGCTCTAT 1205
Db 25728 TGCCCAACACCGAAATCAGATTGACTATATTTCTTTCAGGCAAGAAATGCACTCTAC 25785
Qy 1206 ACAGTCAGCAAAACAGACAGGAGGCTTACTGTGGCTCGAGATCATGACTCTTATGTC 1265
Db 25786 AAAGTCAGCAAAACAGACAGGAGGCTGCTGACTCAGATCATGAACT -CTTATGTC 25844
Qy 1266 CAAATTCAGACTTAAATGAAGAAATAGGAAACCACTAGATCACTCAGTAAAGACCT 1325
Db 25845 CAAATTCAGACTTAAATGAGGAAATAGGAAACCACTAGACCACTCAAGAAATGACCT 25904
Qy 1326 AAATCCAAATCCCTTATGATTATACAGTGGAAATGAGAAATAGATTTAAGGCTTAGATCT 1385
Db 25905 AAATCAAAATCTCTTATGATTATACAGTGGAAATGCAAAATAGATTTAAGGCTTAGATCT 25964
Qy 1386 GATAGCAGAGTACCTAATGAAGAAATGACAGAGGTTTATGACATTTGTACAGGAGACAG 1445
Db 25965 GATAGCAGAGTGCCTAATGAAGTATGACAGAGTTTATGACATTTGTACAGGAGACAG 26024
Qy 1446 GATCGAGACCATCCCATGGAAGAAATGCAAAAGAAAGCAAAATGGCTGCTCTGGGAGGC 1505
Db 26025 GATCAAGACCATCCCATGGAAGAAAGAAAT - - -GCAAAATGGCTGCTCTAAGGAGGC 26076
Qy 1506 CTTTCAAAATAGCTGTGAAGAGAGAGAGTGAAGAGCAAGGAAAGAAAGAAAGATATAA 1565
Db 26077 CTTTCAAAATAGCTGTGAAGAGAGAGAGTCAAAAGCAAGGAAAGAAAGATATAT 26136
Qy 1566 GCATCTGAATGACAGAGTTTCCAAA- - - - - 1588
Db 26137 GCATTTGTTTTCATTCACCCCAAGAAAGCAAGTGCCTCAAGTATATCTCAAACTACCGC 26196
Qy 1589 - - - - - 1588
Db 26197 ACAATTCGACTCATCTCACAGCTAGTAAGTAATGCTCAAAATTTCTCAAGCGAGCTT 26256
Qy 1589 - - - - -GAACTTCAGTGTGTTCAAGCTGGTTTGAAGAAAGTCAAGAGA 1630
Db 26257 CAGCAATATGCGAACCGTGAACCTTCTGATGTTCAAGCTGGTTTGAAGAAAGCAGAGA 26316
Qy 1631 ACCAGAGACCAATTCGCCAAGATCTCTGTATCATATGAGAAAGAAAGAGAGTTCCAGAAA 1690
Db 26317 ACCAGATATCAATTCGCCAATATCCACTGGACCATGGAAGCAAGAGAGTTCCAGAAA 26376
```

QY 1691 AACATCTATTTCTGCTTTTATTGACTATGCAAAAGCCTTTGACTGTGGGGGTCAATAAAA 1750
 Db 26377 AACATCTATTTCTGCTTTTATTGACTATGCCAAGCCTTTGACTGTGGGTCAATAAAA 26436

QY 1751 CTGTGGAAATTCGAAAGGATGGGAATACAGACACACCTGACCTGACTCTTGGAAAAT 1810
 Db 26437 CTGTGGAAATTCGAAAGGATGGGAATACAGACACACCTGACTCTGCTCTTGAGAAC 26496

QY 1811 TTGTATGCAAGTCAGGAAGCAACAGTTAGAACTGCACATGGAACAAACAGACTGCTTCCAA 1870
 Db 26497 TTGTATGTAGTTCAGGAAGCAACAGTTAGAACTGCACATGGAACAAACAGACTGCTTCCAA 26556

QY 1871 GTAGAAAAGAGATGATCAAGCTGTATATGTGCAACCCGCTTGTAACTCTTATGCA 1930
 Db 26557 ATAGAAAAGAGATTCGTCAAGGCTGTATATGTGCAACCCGCTTGTAACTCTTATGCA 26616

QY 1931 GAG-ACATCATGAGAAACGCTGGCTGGAAGACACAGCTGCAATCAAGATTGCCGG 1989
 Db 26617 GAGTACATCATGAGAAACGCTGGCTGGAAGACACAGCTGCAATCAAGATTGCCGG 26676

QY 1990 AGAATACCAATAACCTCAGATATGCAGATATACCAACCTTTATGGCAGAAAGTGAAG 2049
 Db 26677 AGAATATCAATAACCTCAGATATGCAGATATACCAACCTTTATGGCAGAAAGTGAAG 26736

QY 2050 GNACTAAAAGCCTCTTCATGAAGTGAAGAGAGAGTGAAGAGTGGCTTAAAGCTC 2109
 Db 26737 GNACTAAAAGCCTCTTCATGAAGTGAAGAGAGAGTGAAGAGTGGCTTAAAGCTC 26796

QY 2110 AACATTCAAGAAAGCAAGATCATGCGCATCTGGTCCCATCTTCATGGGAAATAGATGG 2169
 Db 26797 AACATTCAAGAAAGCAAGATCATGCGCATCTGGTCCCATCTTCATGGGAAATAGATGG 26856

QY 2170 GAACAGTGGAAACAGTGTGACACTTATTTTGGGGGGCTCCAAATCAGCTGCAGATG 2229
 Db 26857 GAACAGTGGAAACAGTGTGACACTTATTTTGGGGGGCTCCAAATCAGCTGCAGATG 26914

QY 2230 GTGACTGCAGCAGTGAATTAAGAGACACTTACTCTTGGGAAGAAAGTTA--ACCAACC 2287
 Db 26915 GTGACTGCAGCAGTGAATTAAGAGACACTTACTCTTGGGAAGAAAGTTAAGCAACC 26974

QY 2288 TAGTAGATATTAAGAAAGCAGACACTTACTCTTGCACAAAGCCCATCTAGTCAAGG 2347
 Db 26975 TAGTAGATATTAAGAAAGCAGACACTTACTCTTGCACAAAGCCCATCTAGTCAAGG 27034

QY 2348 CTATGGTTTTCCAGTGTGCTATGATGTGAGAGTGGAGTGTGAGAGTGTGAGAGTGTGAGT 2407
 Db 27035 CTATGGTTTTCCAGTGTGCTATGATGTGAGAGTGGAGTGTGAGAGTGTGAGAGTGTGAGT 27094

QY 2408 ACTGAAGAAATTGATGCTTTTGAACGTGGTGTGGGAAGACTCTTGAGAGTCCCTTGG 2467
 Db 27095 GCCAAAGAAATTGATGCTTTTGAACGTGGTGTGGGAAGACTCTTGAGAGTCCCTTGG 27154

QY 2468 CTCGAAGAGATCCAAACAGTCCATCTGAAGAGATCAGCCCTGGGATTTCTTGGAG 2527
 Db 27155 CTCGAAGAGATCCAAACAGTCCATCTGAAGAGATCAGCCCTGGGATTTCTTGGAG 27214

QY 2528 GAATGATGCTAAAGCTGAACTCCAGTACTTTGGCCACCTGATCAGAAGAGCTGACTCAC 2587
 Db 27215 GAATGATGCTAAAGCTGAACTCCAGTACTTTGGCCACCTGATCAGAAGAGCTGACTCAC 27274

QY 2588 TGGAAAAGACCTGATGCTGGAGGATTTGGGGCAGGAGAGAGGGGACGACAGAGGA 2647
 Db 27275 TGGAAAAGACCTGATGCTGGAGGATTTGGGGCAGGAGAGAGGGGACGACAGAGGA 27334

QY 2648 TGAGATGCTGATGGATCACTCACTGATGAGCTGAGTCTGGGTGAACCTCTGAGT 2707
 Db 27335 TGAGATGCTGATGGATCACTCACTGATGAGTCTGAGTCTGAGTGAACCTCTGAGT 27394

QY 2708 TGGTGTAGGACAGGAGGCTCTCTCCGCGCATTCATGGGGTCAAAAGAGTTGACAC 2767
 Db 27395 TGGTGTAGGACAGGAGGCTCTCTCCGCGCATTCATGGGGTCAAAAGAGTTGACAC 27454

QY 2768 GACTGAGCAACTCAACTGAACCTGAACCTG 2795
 Db 27455 GACTGAGCAACTGATCTGATTGATCTG 27482

RESULT 12
 AC144887/c
 LOCUS

DEFINITION
 Bos taurus clone RP42-433E9, WORKING DRAFT SEQUENCE, 8 ordered
 pieces.

ACCESSION
 AC144887
 VERSION
 AC144887.2 GI:31880107
 KEYWORDS
 HTG; HTGS_PHASE2; HTGS_DRAFT.

SOURCE
 Bos taurus
 ORGANISM
 Bos taurus (cow)

REFERENCE
 1 (bases 1 to 196994)
 Antonellis A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K.,
 Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,
 Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,
 Hurle, B., Idol, J.R., Karlins, E., Kwong, P., Laric, P., Lee-Lin, S.-Q.,
 Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C.,
 Maskeri, B., McQuell, J., Peguirigan, C., Pearson, R., Portnoy, M.E.,
 Prasad, A., Reddi-Dugue, N., Schandeler, K., Schueler, M.G., Shah, K.,
 Simon, C., Stantrapop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,
 Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.

AUTHORS

REFERENCE
 2 (bases 1 to 196994)
 Green, E.D.
 Direct Submission
 Submitted (23-MAY-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 196994)
 Green, E.D.
 Direct Submission
 Submitted (18-JUN-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Jun 18, 2003 this sequence version replaced gi:31044307.

TITLE
 NISC Comparative Sequencing Initiative

JOURNAL

REFERENCE
 2 (bases 1 to 196994)

AUTHORS

TITLE
 Direct Submission
 Submitted (23-MAY-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 196994)
 Green, E.D.
 Direct Submission
 Submitted (18-JUN-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Jun 18, 2003 this sequence version replaced gi:31044307.

REFERENCE
 3 (bases 1 to 196994)

AUTHORS

JOURNAL

COMMENT

Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.zoonhgri.nih.gov
 ----- Project Information
 Center project name: eeg
 Center clone name: 433E09

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 195487 bases at least Q40
 Consensus quality: 195921 bases at least Q30
 Consensus quality: 196170 bases at least Q20
 Insert size: 167000; agarose-fp
 Insert size: 196294; sum-of-contigs
 Quality coverage: 12.61x in Q20 bases; agarose-fp


```

ACCESSION      AC139312
VERSION         GI:28460773
KEYWORDS        HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE          Bos taurus (cow)
ORGANISM        Bos taurus

REFERENCE
AUTHORS         Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
                Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
                Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
                Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,B.,
                Leric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
                Marquies,E.H., Masello,C., Maskeri,B., McDowell,J.,
                McGuigan,C., Pearson,R., Portroy,M.E., Prasad,A.,
                Reddi-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
                Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
                Weatherby,K.D., Wiggins,L., Young,A. and Green,E.D.
TITLE           NISC Comparative Sequencing Initiative
JOURNAL         Unpublished
REFERENCE       2 (bases 1 to 192199)
AUTHORS         Green,E.D.
TITLE           Direct Submission
JOURNAL         Submitted (30-JAN-2003) NIH Intramural Sequencing Center, 8717
                Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE       3 (bases 1 to 192199)
AUTHORS         Green,E.D.
TITLE           Direct Submission
JOURNAL         Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
                Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT         On Feb 21, 2003 this sequence version replaced gi:28144407.
                ----- Genome Center
                Center: NIH Intramural Sequencing Center
                Center code: NISC
                Web site: http://www.nisc.nih.gov
                Contact: nisc.zoo@nhgri.nih.gov
                ----- Project Information
                Center project name: ddk
                Center clone name: 482K04

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 190322 bases at least Q40
Consensus quality: 190859 bases at least Q30
Consensus quality: 191112 bases at least Q20
Insert size: 160000; agarose-fp
Insert size: 191299; sum-of-contigs
Quality coverage: 10.93x in Q20 bases; agarose-fp
Quality coverage: 9.14x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and

```

* the accession number will be preserved.

```

* 1 74471: contig of 74471 bp in length
* 74472 74571: gap of unknown length
* 74572 79290: contig of 4719 bp in length
* 79291 79390: gap of unknown length
* 79391 85712: contig of 6322 bp in length
* 85713 85812: gap of unknown length
* 85813 92766: contig of 6954 bp in length
* 92767 92866: gap of unknown length
* 92867 109583: contig of 16717 bp in length
* 109584 109883: gap of unknown length
* 109884 134063: contig of 24379 bp in length
* 134063 134163: gap of unknown length
* 134163 144366: contig of 10204 bp in length
* 144366 144467: gap of unknown length
* 144467 148230: contig of 3764 bp in length
* 148231 148330: gap of unknown length
* 148331 159617: contig of 11287 bp in length
* 159618 159718: gap of unknown length
* 159718 192199: contig of 32482 bp in length.
                Location/Qualifiers
                source
                1..192199
                /organism="Bos taurus"
                /mol_type="genomic DNA"
                /db_xref="taxon:9913"
                /clone="RP42-482K4"
                /clone_lib="RP42"
                1..74471
                /note="assembly_fragment"
                clone_end:SP6
                vector_side:left"
misc_feature
1..86488
/note="clone overlaps with GenBank Accession Number
AC139732 clone RP42-145M1 (center project name ddj)"
misc_feature
74572..79290
/note="assembly_fragment"
79391..85712
/note="assembly_fragment"
/note="assembly_fragment"
85813..92766
/note="assembly_fragment"
92867..109583
/note="assembly_fragment"
106684..192199
/note="clone overlaps with GenBank Accession Number
AC139632 clone RP42-18C2 (center project name ddl)"
misc_feature
109684..134062
/note="assembly_fragment"
134163..144366
/note="assembly_fragment"
144467..148230
/note="assembly_fragment"
148331..159617
/note="assembly_fragment"
159718..192199
/note="assembly_fragment"
clone_end:T7
vector_side:right"

Query Match 58.9%; Score 2105.2; DB 1; Length 192199;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 2581; Conservative 0; Mismatches 153; Indels 608; Gaps 8;

QY 51 GAGCAGTAGTCGGCTTTGCTGCGAGCAGCGGTAAAGAGATACCCACGCCCAAGGTAAGA 110
DB 151713 GAGCATGCTGTGCTTTGCTGCGAGCAGCCATGAAGATACCCACGCCCAAGGTAAGA 151654
QY 111 GAAACCCAGTAAAGATGTTGTGAGAGGCGCATCAGAGGCGACACATCTGAAC 170
DB 151653 GAAATCCAAAGTAAGACGGTAGCTGTTGCAAGAGGCGCATCAGAGGCGCAACACACTGAACC 151594
QY 171 CATACAGCAGGAAACTAGTCAATCTAATCACTAGGACACAGCCTTGCTCTACTCAA 230
DB 151593 CATATCAGAGAAACTAGTCAATCTAATCACTAGGACACAGCCTTGCTGAACTCAA 151534

```

QY 231 TGAAACTAAGCCATGCCCGTGGGCAACCCAAAGATGGCGAGGTCAATGGTGGAGAGATCTG 290
DB 151533 TGAAACTAAGCCATGCCCGTGGGCAACCCAAAGATGGCGAGGTCAATGGTGGAGAGATTTG 151474
QY 291 ACAGAACTGTGTCCACTGGAGAAAGGAAT -GCAAAACCACTTCAGTATCTTTGCTTCGAGA 349
DB 151473 ACAGAACTGTGTCCACTGGAGAAAGGAATGCAAAACCACTTCAGTATCTTTGCTTCGAGA 151414
QY 350 ACCCCATGAAACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGAACTCCCCAGGT 409
DB 151413 ACCCCATGAAACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGAACTCCCCAGGT 151354
QY 410 CAGTAGGTGCCCAATGCTACTCGAGATCAGTGGAGAAATTAATCTCCAGAAAGAACTGAAG 469
DB 151353 CAGTAGGTGCCCAATGCTACTCGAGATCAGTGGAGAAATTAATCTCCAGAAAGAACTGAAG 151294
QY 470 AGATGGAGCCCAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGGTGATATAGCAAGG 529
DB 151293 GGATGGAGCCCAAGCAAAATTAATCTCAGCTGTGGATGTGACTGGTGATAGCAAGG 151234
QY 530 TCCGATGCTTAAAGAGCAATATTGCAATAGGAACCTGGAATGTGAGTCCATGNAATCAAG 589
DB 151233 TCCGATGCTTAAAGAGCAATATTGCAATAGGAACCTGGAATGTGAGTCCATGNAATCAAG 151174
QY 590 GCAAAATGGAAGTGGTCAAAACAGAGATGGCAAGTGAATGTCAAATTTCTAGGAATCA 649
DB 151173 GCAAAATGGAAGTGGTCAAAACAGAGATGGCAAGTGAATGTCAAATTTCTAGGAATCA 151114
QY 650 GCGAACTAAATGGACTGGAATGGTGAATTTAACTCAGATGACCAATATATCTACTACT 709
DB 151113 GTGAACCTAAATGGACTGGAATGGTGAATTTAACTCAGATGACCAATATATCTACTACT 151054
QY 710 GCGGCGAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCGAA 769
DB 151053 GCGGCGAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCTGAA 150994
QY 770 ATGCAGTACTTGGATGCACTCTCAAAACGACAGATGATCTCTGTTTGTTCGAAGGCA 839
DB 150993 ATGCAGTACTTGGATGCACTCTCAAAACGACAGATGATCTCTGTTTGTTCGAAGGCA 150934
QY 830 AACCAATCAATATCACAGTATCAAGTCTATGCCCCAACAGTAACTGCTGAAGAGCTG 889
DB 150933 AACCAATCAATATCACAGTATCAAGTCTATGCCCCAACAGTAACTGCTGAAGAGCTG 150874
QY 890 AAGTTGAACCGTCTATGAAGACCTACAAGACCTTTTAGAACTAAACCCCAAAAGAGATG 949
DB 150873 AAGTTGAACCGGCTATGAAGACCTACAAGACCTTTTAGAACTAAACCCCAAAAGAGATG 150814
QY 950 TCCTTCTCAATATAGGGGACTGGAATGCAAAAGTAGGAGCAAAAGAAACCACTGGAGTAA 1009
DB 150813 TTCTTTTCAATATAGGGGACTGGAATGCAAAAGTAGGAGTCAAGAAACCACTGGAGTAA 150754
QY 1010 CAGGCAAAATTTGGCTTGGATACGGAATGAAGCAGGCGCAAGACTAATAGAGTTTGGC 1069
DB 150753 CAGGCAAAATTTGGCTTGGATATGGAATGAAGCAGGCGCAAGACTAATAGAGTTTGGC 150694
QY 1070 AAGAAATGCACTGGTCTATAGCAAAACCCCTTTTCCAAACCAACAGAGAAAGACTCTACA 1129
DB 150693 AAGAAATGCACTGGTCTATAGCAAAACCCCTTTTCCAAACCAACAGAGAAAGACTCTACA 150634
QY 1130 CATGGACATCACAGAGTGTCAACCGAATCAGATTGATTATTTCTTTGAGGCCAAA 1189
DB 150633 CATGGACATCACAGAGTGTCAACCGAATCAGATTGATTATTTCTTTGAGGCCAAA 150574
QY 1190 GATGGAGAGCTCTATACAGTCAAGCAAAACCAAGCAGGAGCTTACTTGGCTCAGATC 1249
DB 150573 GATGGAGAGCTCTATACAGTCAAGCAAAACCAAGCAGGAGCTTACTTGGCTCAGATC 150514
QY 1250 ATGAACCTCTTATGGCCAAATTCAGACTTAATTTGAAGAGTAGGGAACCACTAGAT 1309
DB 150513 ATGAACCTCTTATTTACCAATTC -GACTTAAATTTGAAGAGTAGGGAACCACTAGAT 150455

QY 1310 CACTCAGGTAAAGACCTAAATCCATCCCTTATGATTATACAGTGAAGTGAAGATAGAT 1369
DB 150454 CACTCAGGTAAAGACCTAAATCCATCCCTTATGATTATACAGTGAAGTGAAGATAGAT 150395
QY 1370 TTAAGGGCCCTAGATCTGATAGACAGAGTACCTTAATGAACCTATGGACAGAGGTTCATGACA 1429
DB 150394 TTAAGGGCCCTAGATCTGAT ---AGAGTGCCTGATGAACCTATGGAAACGAGGTTCATGACA 150339
QY 1430 TTGTACAGAGACAGGGATCGAGACCATCCCATCGAAAGAAATGCAAAAGCAAAAT 1489
DB 150338 TTGTACAGAGACAGGGATCGAGACCATCCCATCGAAAGAAATGCAAAAGCAAAAT 150279
QY 1490 GCTGTCTGGGAGGCTTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAAAGGAA 1549
DB 150278 GCTGTCTGGGAGGCTTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAAAGGAG 150219
QY 1550 AAAAGGAAAGATAAAGAGCATCTGAATGAGAGTTCCTAAA --- 1588
DB 150218 AAAAGGAAAGATAAAGAGCATCTGAATGAGAGTTCCTAAAAGAGATGAAGA 150159
QY 1589 --- 1588
DB 150158 AAGCCTTCTTCTGGCATCAATGCAAAAGAAATAGAGGAAACAAACAGAAATGGAAAGACTA 150099
QY 1589 --- 1588
DB 150098 GGGATCTCTTCAAGAAATCAGAGATACCAAGAGAACTTTTCATGCAAGATGGGCTCGA 150039
QY 1589 --- 1588
DB 150038 TAAAGACAGAAATGGTATGACCTTAAACAGAGACAGAGATATTAAGAGAGATGGCAAG 149979
QY 1589 --- 1588
DB 149978 AATACACAGAGAACTGTACACAAAGATCTTCATGACCCAGATATACCGATGCTGTGA 149919
QY 1589 --- 1588
DB 149918 TCACCTGACCTAGAGCCAGACATCTCTGAAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCA 149859
QY 1589 --- 1588
DB 149858 CTACAAAACAAAGCTAGTGGAGTGAATTCAGATTGAGCTATTCCTCAATCTGAAAG 149799
QY 1589 --- 1588
DB 149798 ATGATGCTGTGAAAGTGTGCACTCAATATGCGCAAAATTTGAAAACTCAGCAGTGGC 149739
QY 1589 --- 1588
DB 149738 CATGGACCTGGAAAAAGGTCAGTTTTCATTTCCAAATCCCAAGAAAGGCAATGCCAAAGAGT 149679
QY 1589 --- 1588
DB 149678 GCTCAAACTACCGTACAAATTGCACTCATCTCACACACAGTAAAGTAATGCTCAAAATTC 149619
QY 1589 ---GAACTTCCAGTTGTTTCAAGCTGGTTTTTA 1616
DB 149618 TTCAAGCCAGGCTTCATCAATATGTGAACCTGTGAACCTTCCAGATGTTTCAAGCTGGTTTTTA 149559
QY 1617 GAAAGTCAAGAGAAACAGAGACCAAAATTTGCCAATCTCTGTATCATGGAAGAAAGCAA 1676
DB 149558 GAAAGTCAAGAGAAACAGAGACCAAAATCAATTTGCCAATCTCTGTGATCGTGGAAAGCAA 149499
QY 1677 GAGAGTTCAGAAAAACATCTATTTCTGCTTTATGCTATGCAAAAGCCTTTGACTGTG 1736
DB 149498 GAGAGTTCAGAAAAACATCTATTTCTGCTTTATGCTATGCAAAAGCCTTTGACTGTG 149439
QY 1737 GGGGTCACAATAAACTGTGGAAAAATTTGAAAGGGATGGGAATACCAACCACTGACCT 1796
DB 149438 TAGATCACAATTAATCTGTGGAAAAATTTGAAAGAGATGGGAATACCAACCACTGACCT 149379
QY 1797 GACTCTTGAAAAATTTGTATGCAGGTGAGGAAGCAACAGTTAGAACTGGACATGGGAACAA 1856


```
Db 149378 GCCTCTTGAGAAATTTGTATGTCAGGTTCAGGAACCAACAGTTGAATGACATGGAACAA 149319
QY 1857 CAGACTGGTTCAGAGTGGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 1916
Db 149318 TAGACTGGTTCAGAGTGGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 149259
QY 1917 TTAACCTTCATGAGAG-ACATCATGAGAAACCTGGCTGGAGAGGAGGAGGAGGAGGAGGAG 1975
Db 149258 TTAACCTTCATGAGAGTGGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 149199
QY 1976 TCAAGATTGCGGAGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 2035
Db 149198 TCAAGATTGCGGAGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 149139
QY 2036 CAGAAAGTGAAGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 2095
Db 149138 CAGAAAGTGAAGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 149079
QY 2096 TTGGCTTAAGCTTCACATTCAGAAACGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 2155
Db 149078 TTGGCTTAAGCTTCACATTCAGAAACGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 149019
QY 2156 GGGAAATAGATGGGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 2215
Db 149018 GGGAAATAGATGGGAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 148961
QY 2216 AATCAGTGCAGATGGTGCAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 2275
Db 148960 AATCAGTGCAGATGGTGCAGGAGTATGTCAAGCGCTGTATTTGTTCACCGGCTGTCT 148901
QY 2276 AGTTA--ACCAACTAGATGAGTATGTCAAGCGGAGGAGTATGTCAAGCGGAGGAGGAGGAGGAG 2333
Db 148900 AGTTAAGCTGAGGAGTATGTCAAGCGGAGGAGTATGTCAAGCGGAGGAGGAGGAGGAGGAG 148841
QY 2334 CCATCTAGTCAAGCTATGTTTTCAGTGGTGCATGTATGTGATGAGATGAGATGAGATGAGATGAG 2393
Db 148840 TCGCTAGTCAAGCTATGTTTTCAGTGGTGCATGTATGTGATGAGATGAGATGAGATGAGATGAG 148781
QY 2394 GAGAAAGCTGAGCAGTGAAGATGATGCTTTGAGTGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 2453
Db 148780 GAGAAAGCTGAGCAGTGAAGATGATGCTTTGAGTGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 148721
QY 2454 GAGAGTCCCTTGGAGTGAAGGAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAG 2513
Db 148720 GAGAGTCCCTTGGAGTGAAGGAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAG 148661
QY 2514 GATTTCTTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 2573
Db 148660 GATTTCTTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 148601
QY 2574 AAGAGCTGAGTCACTGGAAGGAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAG 2633
Db 148600 AAGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 148541
QY 2634 GGGACGACAGAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 2693
Db 148540 GGGACGACAGAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 148481
QY 2694 TGAATCTGAGTGGTGTGAGTGAAGGAGGAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 2753
Db 148480 TGAATCTGAGTGGTGTGAGTGAAGGAGGAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 148421
QY 2754 AAAGAGTGGACAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 2795
Db 148420 AAAGAGTGGACAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAGGAGTGAAG 148383
```

```
RESULT 14
AC142564/c 145088 bp DNA linear \ HTG 02-MAY-2003
LOCUS Bos taurus clone RP42-56B14, WORKING DRAFT SEQUENCE, 2 ordered
DEFINITION pieces.
```

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AC142564
AC142564.2 GI:30315712
HTG; HTGS_PHASE2; HTGS_DRAFT.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 145088)
Antonellis, A., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K.,
Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,
Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,
Huxle, B., Idol, J.R., Karlins, E., Kwong, P., Lalic, P., Lee-Lin, S.-Q.,
Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C.,
Masker, B., McDowell, J., Faguirgan, C., Pearson, R., Portnoy, M.E.,
Prasad, A., Reddi-Dugue, N., Schandler, K., Schueler, M.G., Shah, K.,
Sison, C., Stantripop, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,
Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 145088)
Green, E.D.
Direct Submission
Submitted (04-APR-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 145088)
Green, E.D.
Direct Submission
Submitted (02-MAY-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On May 2, 2003 this sequence version replaced gi:29540599.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: djr
Center clone name: 056B14

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 144975 bases at least Q40
Consensus quality: 144984 bases at least Q30
Consensus quality: 144986 bases at least Q20
Insert size: 159000; agarose-fp
Insert size: 144988; sum-of-contigs
Quality coverage: 14.02x in Q20 bases; agarose-fp
Quality coverage: 15.37x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 32951: contig of 32951 bp in length

* 32952 33051: gap of unknown length

* 33052 145088: contig of 112037 bp in length.

FEATURES

source

1. 145088

1. .145088

```
1. . . . . /organism="Bos taurus"
```

```
/organism="Dros. caudatus"
/mol_type="genomic DNA"
```

```
/db_type="genomic"
/db_xref="taxon:991"
```

```
/usr_xlet="Caxon:991"
/cjone="BP42-56B14"
```

```
/clone="RP42-36B14"
/clone lib="PB42"
```

1 32951

```
L. :32951
/not a "assembly" fragment
```

```
clone and .tr7
/home="assembly
```

```
clone_end:1/  
vector_size:left"
```

```
vector_side:left"
33052 145098
```

33052. .145088
/not a "comb". fragment

```
/note="assembly_in
```

clone_end:SP6

```
vector_side:right'
```

51600. 145088

/note="clone overlaps with GenBank Accession Number

COMMENT On May 23, 2003 this sequence replaced gi:29540597.

```

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoonhgri.nih.gov
----- Project Information
Center project name: djn
Center clone name: 371H23
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 188423 bases at least Q40
Consensus quality: 188607 bases at least Q30
Consensus quality: 188774 bases at least Q20
Insert size: 108000; agarose-fp
Insert size: 189078; sum-of-configs
Quality coverage: 18.32x in Q20 bases; agarose-fp
Quality coverage: 10.45x in Q20 bases; sum-of-configs

```

Query Match	58.8%	Score 2100;	DB 1;	Length 189278;
Best Local Similarity	77.0%;	Pred. No. 0;		
Matches 2617;	Conservative	0;	Mismatches 170;	Indels 610; Gaps 11;
Qy	1	GTCAAGAAGCGCGCAGTGCAGGAGAGGTACTACCTCGTCCAAGTGAAGGACAGTAGC	60	
Db	58236	GTCAAGAAGGTTGGCGGTGAGGAGA-----TACCCCTCATCAAGTGAAGGACAGTGGC	58182	
Qy	61	TGGCTTTGCTGAGCAGACGCGTAA---AGAGATACCCACGCCCAAGGTAGAGAAACCC	117	
Db	58181	TGGCGTTTGTGAGCAGCGTGAAGAGATAGATACCCCATGCCCAAGGTGAGAGAAACCC	58122	
Qy	118	AAGTAAGATCGTAGGTGTTGTGAGAGGGGATCAGAGGGCAGACATCTGAACACCATACAC	177	
Db	58121	AAGTAAGATGTTAGGTGTTTCAAGAGCGGCATCAGAGGGCAACACACATCTGAACCATCTC	58062	
Qy	178	GCAGAAACCTAGTCAATCTAATCAACTAGGACCCAGCGCTGTCTTAATCTCAATGAACCT	237	
Db	58061	ACAGAAACCTAGTCAATCTAATCACAGAGGCCACAGCGCTTGTCTTAATCTCAATGAACCT	58002	
Qy	238	AAGCCATGCCCGCTGGGGCAACCCCAAGATGGCAGGTCATGGTGGAGAGATCTGCACAGAA	297	

Db	58001	AAGCCATGCCCGTGGGCAACCCAAGATGGCGGTCTACTGTGGAGATCTGCAGAAAT	57942
Qy	298	GTGGTCCACTGGAGAAGGAAT-GCAAACCATTTCAGTATTTCTTGCCHTGAGAACCCCAT	356
Db	57941	GTGGTCTACTGGAGAAGAAATGGCAACCACTTCAGTATTTCTTGCCHTGAGAACCCCGT	57882
Qy	357	GAACTATGAAAAGGCAAAATGATAGTACTGAAAGAGGAATCTCCCAGGTCAGTAGG	416
Db	57881	GAACTATGAAAAGGCAAAANGATAGNACTGAAAGGGAATCTCCCAGGTCAGTAGG	57822
Qy	417	TGCCCCATATGTCTGTGGAGATCAGTGGAGAAATAAATCTCCGAAAGAAATGAAGAGATGGA	476
Db	57821	AGCCAGTATGTCTGTGGAGATCAGTAGAGAAATAAATCCAAAAAGAAATGAAGGGATGGA	57762
Qy	477	GCCAAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGTGTGATATAAGCAAGGTCGGATG	536
Db	57761	GCCAAAGCAAAACAATACCCAGCTGTGGATATGACTGTGTGATAGNAGNAGGTCGTATG	57702
Qy	537	CTGTAAAGAGCAATATTCATAGGAACCTTGGAAATCTCAGGTCATTAATCAAGGCAAAAT	596
Db	57701	CTGTAAAGAGCAATATTCATAGGAACCTTGGAAATCTTAGGTCATTAATCAAGGCAAAAT	57642
Qy	597	GGAAGTGTCAACAAGAGATGGCAAGTGAATGT-CAACATTTCTAGGAATCACGGAAC	655
Db	57641	GGAAGTGTCAACAAGAGATGGCAAGTGAATGTCCGACAGTCTAGGAATCAGCAAC	57582
Qy	656	TAAATGGACTGGAATGGGTGAAATTTAACTCAGATGACCAATATATCTACTCTCGGGC	715
Db	57581	TCAAATGGACTGGAATGGGTGAAATTTAACTCAGATGACCAATATATCTACTCTGAGC	57522
Qy	716	AGGAATCCCTCAGAGAAANTGGATAGCCATCATGGTCAAACAAAGAGTCCGAAATGCAG	775
Db	57521	AGGAATCCCTCAGAGAAANTGGGTGGCCATCATGGTCAAC--AAGAGCTCTGAAATGCAG	57464
Qy	776	TACTTGGATCGACTCTCAAAAACGACGAATCATCTGTGTTGTTTCCAAGCAAAACCAT	835
Db	57463	TACTTGGATGCAATCTCAAGAAACAAGAAATCATCTGTGCTTTCCAGGCAAAACCAT	57404
Qy	836	TCAATATCAGTAATCTCAAGTCTATGCCCCAACCAAGTAATGCTGAAGAGCTGAAGTTG	895
Db	57403	TCAATATCAGTATTTCCAACTCTATGCCCCAACCAAGTAATGCTGAAGAGCTGAAGTTG	57344
Qy	896	AAGGTCTTATGAAGCTTACAGACCTTTTGAATCAACAC--CAAAAAGATGTCT	953
Db	57343	AAGGTCTTATGAAGCTTACAGACCTTTTGAATCAACCAAAAAGATGTCT	57284
Qy	954	TCTCATTTAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAACAGG	1013
Db	57283	TTTCATTTAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAACAGG	57224
Qy	1014	CAAAATTTGGCTTTGAAATACGGAAATGAAGCAGGGCAAGACTTAATAGAGTTTTGCCAGA	1073
Db	57223	TAAATTTGGCTTTGGAATACAGAAATGAAGCAGGGCAAAACCTTAATAGAGTTTTGCCAGA	57164
Qy	1074	AAATGCACGTGTCATAGCAAAACACCTCTTTCCAAACAAACAGAGAAGACTCTACACATG	1133
Db	57163	AAATGCACGTGTCATAGCAAAACACCTCTTTCCAAACAAACAGAGAAGACTCTATACATG	57104
Qy	1134	GACATCACCAGATGGTCAACACCGAAATCAGATTAATATATTTCTTTGCAGCCAAAGATG	1193
Db	57103	AACATCACCAGATGGTCAACACCGAAATCAGATTAATATATTTCTTTGCAGCCAAAGATG	57044
Qy	1194	GAGAAGCTCTATACAGTCAGCAAAACAGAGACAGAGGCTTACTGTGCTCAGATCAATGA	1253
Db	57043	GAGAAGCTCTATACAGTCAGCAAAACAGAGACAGAGGCTTACTGTGCTCAGATCAATGA	56984
Qy	1254	ACTCTCTATTGCCAAATTCAGACTTTAAATTTGAAGAAAGTAGGGAACCACTTAGATCACT	1313
Db	56983	ACTCTCTATTACCAATTCAGACTTTAAATTTGAAGAAAGTAGGGAACCACTTAGATCACT	56924
Qy	1314	CAGGTAAGACCTTAATCCAAATCCCTTATGATTTATACAGTGGAGTGAAGAAATAGATTTAA	1373

Db 56923 CAGGTATGACCTAAATCAAAATCCCTTATAATTATACAGTGGAAATGAGAAATAGATTAA 56864
Qy 1374 GGGCTAGATCTGATACAGAGTACCTAATGAACTATGACAGAGTTTCATGACATTTGT 1433
Db 56863 GGGCTAGATCTGATGATAGAGTCCCTGATGAACTAAGGAATGAGTTTCATGACCTTGT 56804
Qy 1434 ACAGGACAGAGGATCGAGACCATCCCTCATGGAAAAGAAATGCAAAAAGCAAAATGGCT 1493
Db 56803 ACAGGATACAGGATCAAGACCATCCCTCATGGAAAAGAAATGCAAAAAGCAAAATGGCT 56744
Qy 1494 GTCTGGGAGCCCTTACAAATAGCTGTGAAAAGAGAGAGTGAAGAAGCAAGCAAAAGAA 1553
Db 56743 GTCTGGGAGGCTTATAATAGCTGTGAAAAGAGAGAGCAAAAG- AAAGGAGAAAA 56685
Qy 1554 GGAAGATAAAGAGCATCTGAATGACAGAGTTCCAAA 1588
Db 56684 GGAAGATAAAGAGCATCTGAAGCAGAGTTCCAAAGATAGCAAGAGATGAAGAAGC 56625
Qy 1589 1589 1588
Db 56624 CTTCTCAGTGATCAATGCAAGAAATAGAGGAAACAAAGAAATGGAAGAGACTGGAGA 56565
Qy 1589 1589 1588
Db 56564 TCTCTCAAGAAATTAGAGATACCAAGGGGACATTTTCATGCAAGATGGGCTTGATAAA 56505
Qy 1589 1589 1588
Db 56504 GGACAGAAATGGTAGCGCCCTAACGGAAGCAGAAGATTAAGAAGAGGTGGCAAGATA 56445
Qy 1589 1589 1588
Db 56444 CATAGAAAGACTGTACAAAAAGATCTTTCAGCAAGCAAGATAATCAAAATGGTGTGATCAC 56385
Qy 1589 1589 1588
Db 56384 TGACCTAGAGCCAGACATCCTGGAAATGTAAGTCAAGTGGTCTTAGGAAGCATCACTAC 56325
Qy 1589 1589 1588
Db 56324 GAAACAAAGCTAGTGGAGTGATGGAAATCCAGCTAAGCTATTTCAAATCCTGAAAGATGA 56265
Qy 1589 1589 1588
Db 56264 TGCTGAGAAAGTGTACACTCAATATGCCAGCAAAATTTGAAAACTCAGCAGTGGCCACA 56205
Qy 1589 1589 1588
Db 56204 GGAATGAAAAGGTCAAGTTTTCATTTCCATCCCAAGAAAGCAATGCCAAAGATGCTC 56145
Qy 1589 1589 1588
Db 56144 AAATCTGACAAATTCGACTCATCTCAGAGTAGTAAAGTAATGTTCAAAATTTCTCCAA 56085
Qy 1589 1589 1623
Db 56084 GCCAGGCTTCAGCAATGTGAACCGTGAACCTCTGATGTTCAAGCTGGTTTGAAGAGG 56025
Qy 1624 CAGAGAACAGAGACCAAAATGGCAACATCTCTGTATCATGGAAAAAGCAAGAGATT 1683
Db 56024 CAGAGAACAGAGATCAAAATGGCAACATCCGCTGGATCATGAAAAAGCAAGAGATT 55965
Qy 1684 CCAGAAAAACATCTATTCTGCTTTTATTTGACTATGCAAAAGCCTTTGACTGTGGGGTCA 1743
Db 55964 CCAGAAAAACATCTATTCTGCTTTTATTTGACTATGCAAAAGCCTTTGACTGTGGATCA 55905
Qy 1744 CAATAAATCTGGAATAATCTGAAAGGATGGAAATACCAAGCACTGACCTGACTCTTT 1803
Db 55904 CAATAAATCTGGAATAATCTGAAAGAGATGGAAATTTCAAGCACTGACCTGCTCTTT 55845
Qy 1804 GAAAAATTTCTATCGAGTCAGGAGCAACAGTTAGAACTGAGATGGACATGGACACACTG 1863
Db 55844 GAAAAATCTGTATGAGGTGAGGAAGCAACAGTTAGAACTGACCTGGAAACACAGACTG 55785

Qy 1864 GTTCAAGTAGGAAAGAGGATATGTCAAGGCTGTATATTTGTCCACCGCTTTTAACTT 1923
Db 55784 GTTCTAATAGGAAAGAGGATACGTCAAGGCTGTATATTTGTCCACCTGCTTATTAACTT 55725
Qy 1924 CTATCCAGAG-ACATCATCAGAAACGCTGGGCTTGAAGAAGCAACAGCTGGAAATCAAGAT 1982
Db 55724 ATATCAGAGTACATCATCAGAAACGCTGGACTGGAAGAAACACAAGCTGGAATCAAGAT 55665
Qy 1983 TGCCCGGAGAAATAGCAATACCTCAGATATGCAGATGATACCAACCTTATGGCAGAAAG 2042
Db 55664 TGCTGGGAGAAATATCAATAACCTCAGATATGCAGATGATACCAACCTTATGGCAGAAAG 55605
Qy 2043 TGAAGAGAACTTAAAAAGCCTCTTCATGAAGGTGAAAGAGAGAGTGAAGAAATTTGGCTT 2102
Db 55604 TGAAGAGAACTTAAAAAGCCTCTTCATGAAGGTGAAAGAGAGCGTGAAGAAATTTGGCTT 55545
Qy 2103 AAAGCTCAACATTCAGAAAAAGAGATCATGAGATCTGGTCCCATCACTTTCATGGGAAAT 2162
Db 55544 AAAGCTCAACATTCAGAAAAAGAGATCATGAGATCTGGTCCCATCACTTTCATGGGAAAT 55485
Qy 2163 AGATGGGAAACAGTGGAAACAGTGTGAGCTTTATTTTGGGGGGCTTCCAAATCACT 2222
Db 55484 AGATGGGAAACAGTGGAAACAGTGTGAGCTTTA- TTTTGGGGGGCTTCAAAATCACT 55427
Qy 2223 GCAGATGTGACTGCAGCCATGAAATTAAGACACTTACTCTCTTGGAGAAAGATTA- 2280
Db 55426 ATGATGTGACTGCAGCCATGAAATTAAGACACTTACTCTCTTGGAGAAAGATTTATG 55367
Qy 2281 ACCAACCTTAGATAGCATATTGAAAGAGAGACATTTACTTTGCCAACAAAGCCCATCTA 2340
Db 55366 ACCAACCTTAGATAGCATATTCCAAAGCAGACATTTACTTTGCCAACAAAGGTTCTGCTA 55307
Qy 2341 GTCAAGGCTATGGTTTTTCCAGTGTGATGATGATGAGTGTGAGTGTGAGTGTGAGAAA 2400
Db 55306 GTCAAGGCTATGGTTTTTCCAGTGTGATGATGATGAGTGTGAGTGTGAGTGTGAGAAAG 55247
Qy 2401 GCTGAGCACTGAGAAATGATGCTTTTGAACCTGGTGTGAGAGAGACTCTTGAAGTCT 2460
Db 55246 GCTGAGCCCAAGAAATTTGATGCTTTTGAACCTGGTGTGAGAGACTCTTGAAGTCT 55187
Qy 2461 CTTTGCAGCTGCAAGGAGATCCAAACAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCT 2520
Db 55186 CTTTGCAGCTGCAAGGAGATCCAAACAGTCCATTTCTGAAGGAGATCAGCCCTGGGATTTCT 55127
Qy 2521 TTGGAAGGAATGATCTCTAAAGCTGAAATCCAGTATCTTTGGCCACTGATCAGAAAGACT 2580
Db 55126 TTGGAAGGAATGATCTCTAAAGCTGAAATCCAGTATCTTTGGCCACTGATCAGAAAGACT 55067
Qy 2581 GACTCACTGGAAGAACCCCTGATGCTGGAGGAGTGGGGCAGAGAGAGAGGGGAGCA 2640
Db 55066 GACTCACTGGAAGAACCTGATGCTGGAGGAGTGGGGCAGAGAGAGAGAGGGGAGCA 55007
Qy 2641 CAGAGATCAGATGGCTGGATGATCACTGACTGATGAGTGTGAGTGTGGTGAATCT 2700
Db 55006 CAGAGATCAGATGGCTGGATGATCACTGACTGATGAGTGTGAGTGTGAGTGTGAGTGT 54947
Qy 2701 CTGGATTTGGTATGGAAGAGAGGCTCTCTGGGGGATTCATGGGTCACAAAGAGT 2760
Db 54946 CGGGATTTGGTATGGAAGAGAGGCTCTGGATGCTGCGATTCATGGGGCTGTGAGAGT 54887
Qy 2761 TGAACCACTCAGCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAA 2797
Db 54886 CGGACACCACTCAGGAGCTGAAAGGAACTGAACTGAACTGAACTGAACTGAACTGAA 54850

RESULT 16

AC139632/c

LOCUS

DEFINITION

AC139632

ACCESSION

VERSION

AC139632
Bos taurus clone RP42-18C2, WORKING DRAFT SEQUENCE, 9 ordered
pieces
AC139632
GI:29150353

224721 bp DNA linear HTG 22-MAR-2003

KEYWORDS
SOURCE
ORGANISM

HTG; HTGS_PHASE2; HTGS_DRAFT.
Bos taurus
Bos taurus (cow)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS

1 (bases 1 to 224721)
Akhter, N., Antonellis, A., Ayle, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
Laric, P., Lee-Ihn, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Marquies, E.H., Masello, C., Maskeri, B., McDowell, J.,
Paguigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,
Redix-Bugue, N., Schandler, K., Schueler, M.G., Sison, C.,
Santripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative

TITLE
JOURNAL

Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 224721)
Green, E.D.

TITLE
JOURNAL

Direct Submission
Submitted (07-FEB-2003) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA

REFERENCE
AUTHORS

3 (bases 1 to 224721)
Green, E.D.

TITLE
JOURNAL

Submitted (22-MAR-2003) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
On Mar 22, 2003 this sequence version replaced gi:28269383.

COMMENT

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoehghri.nih.gov
----- Project Information
Center project name: ddi
Center clone name: 018C02

The sequence data in this record represents an 'enhanced' version of a phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 222839 bases at least Q40
Consensus quality: 223479 bases at least Q30
Consensus quality: 223749 bases at least Q20
Insert size: 191000; agarose-fp
Insert size: 223921; sum-of-contigs
Quality coverage: 11.77x in Q20 bases; agarose-fp
Quality coverage: 10.04x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 2461: contig of 2461 bp in length

* 2462	2561: gap of unknown length	
* 2562	42437: contig of 39876 bp in length	
* 42438	42537: gap of unknown length	
* 145666	contig of 103129 bp in length	
* 145667	gap of unknown length	
* 147112	contig of 1346 bp in length	
* 147113	gap of unknown length	
* 147113	contig of 33381 bp in length	
* 180593	contig of 33381 bp in length	
* 180594	gap of unknown length	
* 185471	contig of 4778 bp in length	
* 185472	gap of unknown length	
* 185572	contig of 2269 bp in length	
* 187840	contig of 2269 bp in length	
* 187841	gap of unknown length	
* 215656	contig of 27716 bp in length	
* 215657	gap of unknown length	
* 215757	contig of 8965 bp in length.	
FEATURES		
source	1..224721	
	/organism="Bos taurus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9913"	
	/clone="RP42-18C2"	
	/clone_lib="RP42"	
misc_feature	1..88516	
	/note="clone overlaps with GenBank Accession Number AC139312 clone RP42-482K4 (center project name ddk)"	
misc_feature	1..2461	
	/note="assembly_fragment"	
	clone_end:SP6	
	vector_side:left"	
misc_feature	2562..42437	
	/note="assembly_fragment"	
misc_feature	42538..145666	
	/note="assembly_fragment"	
misc_feature	145767..147112	
	/note="assembly_fragment"	
misc_feature	147213..180593	
	/note="assembly_fragment"	
misc_feature	180594..185471	
	/note="assembly_fragment"	
misc_feature	185572..187840	
	/note="assembly_fragment"	
misc_feature	187841..215656	
	/note="assembly_fragment"	
misc_feature	200093..224721	
	/note="clone overlaps with GenBank Accession Number AC139311 clone RP42-242D1 (center project name ddm)"	
misc_feature	215757..224721	
	/note="assembly_fragment"	
	clone_end:T7	
	vector_side:right"	
Query Match 58.7%; Score 2095.4; DB 1; Length 224721;		
Best Local Similarity 77.2%; Pred. NO. 0;		
Matches 2561; Conservative 0; Mismatches 151; Indels 610; Gaps 9;		
QY	51	GAGCAGTACGCGCTTTGCTGGACGACCGTAAAGAGATACCCACGCCCAAGTAAGA 110
Db	47119	GAGCAATGGCTGTGCTTTGCTGGACGACCGTAAAGAGATACCCACGCCCAAGTAAGA 47060
QY	111	GAACCCCAAGTAAGTAGTGTGTGTGAGAGGCGCATCAGAGGCGACATACCTGAAC 170
Db	47059	GAATCCCAAGTAAGTAGTGTGTGTGCAAGAGGCGCATCAGAGGCGACATACCTGAAC 47000
QY	171	CATACACGCGAAGAACTAGTCAATCACTACACTAGGACACACCGCTTGTCTACTCAA 230
Db	46999	CATATCTCAGAAAACCTAGTCAATCACTACACTAGGACACACCGCTTGTCTACTCAA 46940
QY	231	TGAACCTAAGCCATGCCGTGGGCAACCCCAAGATGGCAGGTCTATGTTGGAGAGATCTG 290
Db	46939	TGAACCTAAGCCATGCCGTGGGCAACCTAAGAGGCGCAGGTCTATGTTGGAGAGATCTG 46880
QY	291	ACAGAAATGTGTTCCACTCGAGAGAGGAAT-GCAACACCACTTCAGTATTCTTGCCTTGAGA 349

Db 46879 ACAGAAATGTGGTCCACTGAGAGGAATGCAAAACCACTTCAGTATTTCTTGGCTTGGA 46820
QY 350 ACCCATGAACAGTATGAAAGGCAAAATGATAGTACTGAAAGAGAACTCCCCAGGT 409
Db 46819 ACCCATGAACAGTATGAAAGGCAAAATGATAGTACTGAAAGAGAACTCCCCAGGT 46760
QY 410 CAGTAGTGGCCCCATATCTACTGAGATCAGTGGAGAAATAAATCCAGAAAGAAATGAAG 469
Db 46759 CAGTAGTGGCCCCATATCTACTGAGATCAGTGGAGAAATAAATCCAGAAAGAAATGAAG 46700
QY 470 AGATGGAGCCAAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGTGGTATATAAGCAAGG 529
Db 46699 GGATGGAGCCAAAGCAAAAGAAATAAATCTCAGTGTGGATGTGACTGGTATAGAAAGCAAGG 46640
QY 530 TCCGATGCTGTAAGAGCAAAATATGCTAGGAACCTGGAAATGTGAGTCCATGAATCAAG 589
Db 46639 TCCGATGCTATAGAGGCAAAATATGCTAGGAACCTGGAAATGTGAGTCCATGAATCAAG 46580
QY 590 GCAAAATGGAGTGGTCAAAACAAGAGATGGCAAGTGAATGTCAACATTTAGGAATCA 649
Db 46579 GCAAAATGGAGTGGTCAAAACAAGAGATGGCAAGTGAATGTCAACATTTAGGAATCA 46522
QY 650 GCGAATCAATATGAGTGGTCAAAACAAGAGATGGCAAGTGAATGTCAACATTTAGGAATCA 709
Db 46521 GTGAATCAATATGAGTGGTCAAAACAAGAGATGGCAAGTGAATGTCAACATTTAGGAATCA 46462
QY 710 GCGGAGAGAAATCCCTCAGAAAGAAATGGAGTGGTCAATCATGGTCAACAAAAGAGTCCGAA 769
Db 46461 GCGGAGAGAAATCCCTCAGAAAGAAATGGAGTGGTCAATCATGGTCAACAAAAGAGTCCGAA 46402
QY 770 ATGCAGTACTTGGATGCAATCTCAAAACGACAGAAATGATCTGTGTTGTTTCAAGGCA 829
Db 46401 ATGCAGTACTTGGATGCAATCTCAAAACGACAGAAATGATCTGTGTTGTTTCAAGGCA 46342
QY 830 AACCATTCATATCACAGTATCAAGTCTATGCCCCAACAGTAATGCTGAAGAGCTG 889
Db 46341 AACCATTCATATCACAGTATCAAGTCTATGCCCCAACAGTAATGCTGAAGAGCTG 46282
QY 890 AAGTTGAACGGTCTATGAAGACCTCAAGACCTTTTAGAACCTAACCCCAAAAGAGATG 949
Db 46281 AAGTTGAACGGTCTATGAAGACCTCAAGACCTTTTAGAACCTAACCCCAAAAGAGATG 46222
QY 950 TCCCTTCTATATAGGGGAGCTGGAATGCAAAAGTAGGAGCAAGCAACCTGGAGTAA 1009
Db 46221 TCCCTTCTATATAGGGGAGCTGGAATGCAAAAGTAGGAGTCAAGAACCAACCTGGAGTAA 46162
QY 1010 CAGCAAAATTTGGCTTTGGATACGGAATGAAGCAGGGCAAGAACTAATAGATTTTGCC 1069
Db 46161 CAGCAAAATTTGGCTTTGGATATGGAATGAAGCAGGGCAAGAACTAATAGATTTTGAC 46102
QY 1070 AAGAAATGCACTGGTATAGCAACACCTCTTCCAAACACACAGAGAGAGCTTACA 1129
Db 46101 AAGAAATGCACTGGTATAGCAACACCTCTTCCAAACACACAGAGAGAGCTTACA 46042
QY 1130 CATGGACATACCAAGATGGTCAACACCGAAATCAGATTTGATTTATTTTGGAGCCAA 1189
Db 46041 CATGGACATACCAAGATGGTCAACACCGAAATCAGATTTGATTTATTTTGGAGCCAA 45982
QY 1190 GATGGAGAGCTTATACAGTACGCAAAACACAGCAGAGCTTACTGTGGTCAAGTC 1249
Db 45981 GATGGAGAGCTTATACAGTACGCAAAACACAGCAGAGCTTACTGTGGTCAAGTC 45922
QY 1250 ATGAACCTCTTTTGGCAAAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGAT 1309
Db 45921 GTGAACTCTTTTACCAAAATTC-GACTTAAATTTGAAGAAAGTAGGAAACCACTAGAT 45863
QY 1310 CACTCAGGTAAAGCTTAAATTCCTTATGATTTATACAGTGGAAAGTAGGAAATAGAT 1369
Db 45862 CACTCAGGTAAAGCTTAAATTCCTTATGATTTATACAGTGGAAAGTAGGAAATAGAT 45803
QY 1370 TTAGGCGCTAGTCTGATAGACAGTACTTAATGAATATGACAGAGGTTTCATGACA 1429

Db 45802 TTAAGGGCTTAGATCTTGAT-----AGATGCCTGATGAACCTATGGAACGAGGTTTCATGACA 45747
QY 1430 TTGTACAGGAGACAGGGATCGAGACCATCCCATGGAAGAAATGCAAAAGCAAAAT 1489
Db 45746 TTGTACAGGAGACAGGGATCGAGACCATCCCATGGAAGAAATGCAAAAGCAAAAT 45687
QY 1490 GGCTGTCTGGGAGGCTTACAAATAGCTGTGAAAAAGAGAGAGTGAAGAGCAAGGAA 1549
Db 45686 GGCTGTCTGGGAGGCTTACAAATAGCTGTGAAAAAGAGAGAGTGAAGAGCAAGGAG 45627
QY 1550 AAAAGAAAGATAAAAAGCATCTGAATGCAAGTTCCTCAAA----- 1588
Db 45626 AAAAGAAAGATATAAACATCTGAATGCAAGTTCCTCAAAAGATAGCAAGAGAGATAAGA 45567
QY 1589 ----- 1588
Db 45566 AAGCCTTCTTCTCGATCAATGCAAAAGAAATAGAGGAAACCAACAGATGGGAAAGACTA 45507
QY 1589 ----- 1588
Db 45506 GGGATCTCTTCAAGAAATACAGAGATACCAAGGAACATTTTCATGCAAGATGGGCTCGA 45447
QY 1589 ----- 1588
Db 45446 TAAAGACAGAAATGCTATGGAACCTTAACAGAGAGCAAGATATTAAGAAAGATGGAAG 45387
QY 1589 ----- 1588
Db 45386 AATACACAGAGAACTGTATACAAAGAGATCTTCATGACCCAGATATACAGATGGTGTGA 45327
QY 1589 ----- 1588
Db 45326 TCACTGACCTAGAGCCAGACATCTCTGGAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCA 45267
QY 1589 ----- 1588
Db 45266 CTACAAACAAAGCTAGTGGAGGTGATGGAATTCACAGTTGAGCTATTCCTCAAAATCCTGGAAG 45207
QY 1589 ----- 1588
Db 45206 ATGATGCTGTGAAAGTGTGCTCACTCAATATGCCAGCAAAATTTGGAAAACTCAGCAGTGGC 45147
QY 1589 ----- 1588
Db 45146 CATGGACTGGAAGAGGTGAGTCACTTTTCATTCCTCAATCCCAAGAAAGCAATGCCAAGAGT 45087
QY 1589 ----- 1588
Db 45086 GCTCAAACTACCGTACAAATTCGACTCACTCACTCACACAGTAAAGTAAATGCTCAAAATTC 45027
QY 1589 -----GAACTTCCAGTTGTTCAAGCTGGTTTGA 1616
Db 45026 TTCAAGCCAGGCTTCATCAATATGTGAATGTGAATGTGAATGTTCAGATGTTCAAGCTGGTTTGA 44967
QY 1617 GAAAGTACAGAGAACCCAGAGACCAAAATTTGCCAATCTCTGTATCATGGAAGAAAGCA 1676
Db 44966 GAAAGTACAGAGAACCCAGAGATCAAAATTTGCCAATCTCTGTGTGATCGTGGAAAGCA 44907
QY 1677 GAGAGTTCCAGAAAAACATCTATTTCTGCTTTTACTATGCTATGCTTCAAAAGCTTTGACTGTG 1736
Db 44906 GAGAGTTCCAGAAAAACATCTATTTCTGCTTTTATGCTATGCTTCAAAAGCTTTGACTGTG 44847
QY 1737 GGGGTACAAATAAATCTGGAAGAAATTTCTGAAGGATGGGAATACAGACCACTGACCT 1796
Db 44846 TGGATACAAATAAATCTGGAAGAAATTTCTGAAGAGATGGGAATACAGACCACTGACCT 44787
QY 1797 GACTCTTGAAGAAATTTGTATGAGGTACAGAGCAACAGTTAGAACTGACATGGAACAA 1856
Db 44786 GCCTCTTGAGAAATTTGTATGAGGTACAGAGCAACAGTTAGAACTGACATGGAACAA 44727
QY 1857 CAGACTGGTTTCAAGTAGGAAAGAGATGTCAAGGCTGTATATTTGACCCGCTTGT 1916
Db 44726 TAGACTGGTTTCAAAATAGGAAAGAGTTCATCAAGGCTGTATATTTGTCACCCCTGTTAT 44667

```
Query Match          58.3%; Score 2081.09; DB 1; Length 183383;
Best Local Similarity 75.5%; Pred. No. 0;
```

		Matches 2669; Conservative 0; Mismatches 244; Indels 623; Gaps 13;	
Qy	6	GAAGGGCGGAGTGAAGAGAGTACCTACCTCGTCCAGGTAAAGAGAGAGTGTGCGC	65
Db	69888	GAAGGGCGGAGTGAAGAGAGTACCTACCTCGTCCAGGTAAAGAGAGAGTGTGCGC	69942
Qy	66	TTTGTCTGAGCAGCGTAAAGAGATACCCACGCCCCAAGTGAAGAGAAACCCCAAGTAAAG	125
Db	69943	TTTGTCTGAGCAGCGTGAAGAGATACCCACGCCCCAAGTGAAGAGAAACCCCAAGTAAAG	70002
Qy	126	TGTTAGTGTGTGAGAGGCGATCAGAGGCGAGACATCTGAAACCAATACACGCAAAAA	185
Db	70003	CGGTAGTGTGTGAGAGGCGATCAGAGGCGAGACATCTGAAACCAATACACGCAAAAA	70061
Qy	186	CTAGTCAATCTAATCAACATAGACACACGCTTGTCTAATCAATGAACCTAAGCCATG	245
Db	70062	CTAGTCAATCTAATCAACATAGACACACGCTTGTCTAATCAATGAACCTAAGCTATG	70121
Qy	246	CCCCTGGGGCAACCCAGATGGGAGGTCTATGGTGGAGAGATCTGACAGATGTGTGCTCA	305
Db	70122	CCCCTGGGGCAACCCAGATTTGGTGGGTCTATGGTGGAGAGATCTGACAGATGTGTGCTCA	70181
Qy	306	CTGGAGAGAGGAAT-GCAAAACCACTTCAGTATCTTGTGCTTGAACCCCAATGAACAGTA	364
Db	70182	CTGGAGAGAGGAATGGCAAAACCACTTCAGTATCTTGTGCTTGAACCCCAATGAACAGTA	70241
Qy	365	TGAAAAGGGCAAAATGATAGGATCTGAAAGAGAACTCCCCAGGTCACTAGGTGCCCCAT	424
Db	70242	TGAAAAGGGCAAAATGATATGATCTGAAAGAGAACTCCCCAGGTCACTAGGTGCCCTAAT	70301
Qy	425	ATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCCAAAGC	484
Db	70302	ATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAGTGAAGGATGGAGCCAAAGC	70361
Qy	485	AAAAAGAAATCCAGCTGTGGATGTGATGCTGATATAGCAAGGTCCGATGCTGTAAG	544
Db	70362	AAAAAGAAATCCAGCTGTGGATGTGATGCTGATATAGCAAGGTCCGATGCTGTAAG	70421
Qy	545	AGCAATATTGCATAGGAACCTGGAATGTCAGGTCCATGATCAAGGCAAAATGGAAGTGG	604
Db	70422	AGCAATATTGCATAGGAACCTGGAATGTCAGGTCCATGATCAAGGCAAAATGGAAGTGG	70481
Qy	605	TCAAAAGAGATGGCAAGAGTGAATGTCAACATCTTAGGAATCAGCGCAACTAAATGGA	664
Db	70482	TCAAAAGAGATGGCAAGAGTGAATGTGACATCTTAGGAATCAGTGAACCTGAAATGGA	70541
Qy	665	CTGGATGGGTGAATTTAACTCAGATGACCAATATATCTACTACTGCGGCGAGGAATCCC	724
Db	70542	CTGG-ATGGGTGACTTTAACTCAGATGACCAATATATCTACTACTGCGGCGAGGAATCCC	70600
Qy	725	TCAGAGAAATGGAGTAGCCATCATGTGTCAACAAAAGATCCGAAATGCAAGTACTTGGAT	784
Db	70601	TCAGAGAAATGGAGTGGCCATCATGTGTCAACAAAAGATCTGAAATACAGTACTTGGAT	70660
Qy	785	GCGTCTCAAAAACGACAGATGATCTCTGTTGTTTCCAAAGGCAACCAATTCATATCA	844
Db	70661	GCAATTTCAAAAACGACAGATGATCTCTGTTGTTTCCAAAGGCAACCAATTCATATCA	70720
Qy	845	CAGTAAATCCAGTCTATGCCCCCAACCAAGTAACTGGAAGAGCTGAAGTTGAACGGTCTCT	904
Db	70721	CAGTAAATCCAGTCTATGCCCCCAACCAAGTAACTGGAAGAGCTGAAGTTGAACGGTCTCT	70780
Qy	905	ATGAGACCTTACAGACCTTTTAGAATCAACCCC-AAAAAGATGTCCTTCTCATATTA	963
Db	70781	ATGAGACCTTACAGACCTTTTAGAATCAACCCC-AAAAAGATGTCCTTCTCATATTA	70840
Qy	964	GGGAGCTTGGAAATGCAAAAGTAGGAAGCAAAAGAAACACCTGGAGTAAACAGGCAAAATTTGGC	1023
Db	70841	GGGAGCTTGGAAATGCAAAATAGGAAGTCAAGAAACACCTGGAGTAAAGGCAAAATTTGGC	70900
Qy	1024	CTTGGAAATAGCGAATGAAGAGGCGCAAGACTAATAGAGTTTGGCCAGAAATATGCACTG	1083
Db	70901	CTTGGAAATAGCGAATGAAGAGGCGCAAGACTAATAGAGTTTGGCCAGAAATATGCACTG	70960

Qy	1084	GTCCATAGCAAAACACCTCTTCCAAACAACACAGAGAAAGACTTACACATGACATCACCA	1143
Db	70961	GTCCATAGCAAAACACCTCTTCCAAACAACACAGAGAAAGACTTATACATGACATCACCA	71020
Qy	1144	GATGTGTCAACCCGAAATCAGATTGATTATATCTTTTCAGCCAAAGATGGAGAAAGCTCT	1203
Db	71021	GATGTGTCAACCCGAAATCTGATTGATTATATCTTTTCAGCCAAAGATGGAGAAAGCTCT	71080
Qy	1204	ATACAGTCAGCAAAAACCAAGACCCAGGAGCTTACTGTGGCTCAGATCATCTCCTTATTT	1263
Db	71081	ATACAGTCAGCAAAAACCAAGACCCAGGAGCTGACTGTGGCTCAGACCATGAACTCCTTATTT	71140
Qy	1264	GGCAAAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAAAGAC	1323
Db	71141	GGCAAAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGACCATTCAGGTATGAC	71200
Qy	1324	CTAAATCCAAATCCCTTATGATTATACAGTGAAGTGAAGAAATAGATTTTAAAGGCTAGAT	1383
Db	71201	CTAAATCCAAATCCCTTATGATTATACAGTGAAGTGAAGAAATAGATTTTAAAGGCTAGAT	71260
Qy	1384	CTGATAGACAGATACCTTAATGAACCTATGGACAGAGGTTCATGACATTTGTACAGAGACA	1443
Db	71261	CTGATAGATAGAGTGCCTGATGAACCTATGGAAATGAGGTTCATGACATTTGTACAGAGACA	71320
Qy	1444	GGGATCGAGACCATCCCATGGAAAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGGAG	1503
Db	71321	GGGATCAAGACCATCCCATGGAAAAGAAATGCAAAAAGCGAAATGCTGTCTGGGGAG	71380
Qy	1504	GGCTTACAATAGCTGTGAAAAGAGAGAGTGAAGAAACCAAGCAAAAGGAAAGATAA	1563
Db	71381	GGCTTACAATAGCTGTGAAAAGAGAGAGTGAAGAAACCAAGCAAAAGGAAAGATAA	71437
Qy	1564	AAGCATCTGAATGACAGAGTTCCTCAAA-	1588
Db	71438	AAACATCTGAATGACAGAGTTCCTCAAAAGATAGCAAGAGATAGAAAGCTTCTTCAGC	71497
Qy	1589	-----	1588
Db	71498	GATCAATGCAAAAGAAATAGAGGAAACCAACAGATGGGAAAGACTAGGGATCTCTTCAAG	71557
Qy	1589	-----	1588
Db	71558	AAATTCAGAGATACCAAAAGGAACATTTCTATGCAAAAGATGGGCTCAATTAAGGACAGAAAT	71617
Qy	1589	-----	1588
Db	71618	GGTATGGACCTAAACAGAGCAGAGATATTAAGAGACATGGCAATAATACACAGAGAA	71677
Qy	1589	-----	1588
Db	71678	CTGTACAGAAAAGATCTTTCATGACCCAGATAATCAATGGTGTGATCACTGACCTAGAG	71737
Qy	1589	-----	1588
Db	71738	CCAGACATCTCGAATGTGAAGTCAAGTGGGCTTAGAAAGCATCACTACAAACAAAGCT	71797
Qy	1589	-----	1588
Db	71798	AGTGAGGTGATGGAATTCAGATTGAGCTATTCCAAATCTCTGAAAGATGATGCTGTGAAA	71857
Qy	1589	-----	1588
Db	71858	GTGTTGCACTCAATATGCCAGCACATTTGGAAAACCTCAGCAGTGGCCACAGGACTGGAAA	71917
Qy	1589	-----	1588
Db	71918	AGGTCAGTTTTTCATTCCCAATCCCAAGAAAGCAATGCCAAAGATGCTCAAACTACTGC	71977
Qy	1589	-----	1588
Db	71978	ACAATTCACCTCATCTCACGCTAGTAAAGTAAATGCTCAAAATTTCTCAAGCCAGGCTT	72037

QY 1589 -----GAACTCCAGTGTGTTCAAGCTGGTTTTGAAGAAGTCAGAGG 1630
 Db 72038 CAGCAATATGTGAACCATGAAATTCCTGATGTTCAAGCTGGTTTTGAAGAAGTCAGAGG 72097
 QY 1631 ACCAGAGACCAAAATGTGCAACATCCTCTGTATCATGTGAAAAAGCAAGAGAGTTCAGAGAA 1690
 Db 72098 ACCAGAGATCAATATGGCCCAACATCCGCTGGATCATGGAAGAAAGCAAGAGAG-TCCAGAAA 72156
 QY 1691 AACATCTATTTCTGCTTATGTACTATGCAATGCAAAAGCCCTTCACTGTGGGGTCTCAATATA 1750
 Db 72157 AGCATCTATTTCTGCTTATGTACTATGCAATGCAAAAGCCCTTCACTGTGGGGTCAAGAAA 72216
 QY 1751 CTGTGGAATAATCTGAAAGGATGGGAATACCAAGACCACTGACCTGACCTTCCTTGAAGAAAT 1810
 Db 72217 CTGTGGAATAATCTTCAAGAGATGGGAATACCAAGACCACTGACCTGACCTTCCTTGAAGAAAT 72276
 QY 1811 TTGTATGAGTTCAGGAAGCAAGTGTAGAACTGGAACATGGAACCAAGACTGGTTCCTCAA 1870
 Db 72277 TTGTATGAGTTCAGGAAGCAAGTGTAGAACTGGAACATGGAACCAAGACTGGTTCCTCAA 72336
 QY 1871 GTAGSAAAAGGAGTATGTCAAGGCTGTATATGTCAACCGGCTTGTCTTAACCTTCTATGCA 1930
 Db 72337 ATAGSAAAAGGAGTATGTCAAGGCTGTATATGTCAACCGGCTTGTCTTAACCTTCTATGCA 72396
 QY 1931 GAG-ACATCATGAGAAACGCTGGCTGGAGAGAACACAAAGCTGGAATCAAGATTCGCGGG 1989
 Db 72397 GAGTACATCATGAGAAACGCTGGCTGGAGAGAACACAAAGCTGGAATCAAGATTCGCGGG 72456
 QY 1990 AGAATAGCAATACCTCAGATATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2049
 Db 72457 AGAATATCAATACCTCAGATATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 72516
 QY 2050 GAACTAAAAGCCCTCTTGATGAAGCTGAAAGAGAGAGAGTGAAGAGTGGCTTAAAGCTC 2109
 Db 72517 GAACTAAAAGCCCTCTTGATGAAGCTGAAAGAGAGAGTGAAGAGTGGCTTAAAGCTC 72576
 QY 2110 AACATTCAGAAAAGCAAGATCATGGCATCTGTGTCCTCATCTTCATGCGGAAATAGATGGG 2169
 Db 72577 AACATTCAGAAAAGCAAGATCATGGCATCTGTGTCCTCATCTTCATGCGGAAATAGATGGG 72635
 QY 2170 GAAACAGTGGAAACAGTGTGAGACTTATTTTGTGGGGGCTCAAAATCACTGCAAGATG 2229
 Db 72636 -----GGAACAGTGTGAGACTTATTTT--CTGGGCTCAAAATCACTACAGATG 72685
 QY 2230 GTGACTGAGCCATGAAATTAAGAGACACTTCTCTTGGAGAGAAAGTTA--ACCAACC 2287
 Db 72686 GTGACTTCAAGCATGAAATTAAGAGAGCGCTTACTCTTGGAGAGAAAGTTATGACCAACC 72745
 QY 2288 TAGATGACATATTCAGAAAGAGAGACATTAACCTTGCACCAAGAGCCCATCTAGTCAAGG 2347
 Db 72746 TAGATGACATATTCAGAAAGAGAGACATTAACCTTGCACCAAGAGCGCTTCACTAGTCAAGG 72805
 QY 2348 CTATGGTGTTCCTAGTGTGATGTGAGAGTGTGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 2407
 Db 72806 CTATGGTGTTCCTAGTGTGATGTGAGAGTGTGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 72862
 QY 2408 ACTGAGAAATGTGCTTTGAACTGTGGTGTGGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 2467
 Db 72863 GCTAAAGAAATGTGCTTTGAACTGTGGTGTGGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 72922
 QY 2468 CTGCAAGGAGATCAACACAGTCCATCTGAGAGAGATCAGCCCTGGGATTTCTTTTGAAG 2527
 Db 72923 CTGCAAGGAGATCAACACAGTCCATCTGAGAGAGATCAGCCCTGGGATTTCTTTTGAAG 72982
 QY 2528 GAATGATGCTAAAGCTGAAATCTCAGTACTTTTGGCCACCTGATCAGAAAGAGTGTGACTCAC 2587
 Db 72983 GAATGATGCTAAAGCTGAAATCTCAGTACTTTTGGCCACCTCATGCAAAAGAGTGTGACTCAT 73042
 QY 2588 TGGAAAAGACCTCATGCTGGGGGATTTGGGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2647
 Db 73043 TGGAAAAGACCTCATGCTGGAGAGGATTTGGGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 73102
 QY 2648 TGAGATGGCTGAGTGGCATCTGACTCGATGGAGAGTGTGCTGGGTGAACCTCTCTGGAGT 2707

Db 73103 TGAGATGCCGAGTGCATCACTGACTCAGTGGATGTGAGTCTGGTGAATTCGGGAGT 73162
 QY 2708 TGGTGTATGAGACAGGAGGCGCTGCTCGCGGAGTTCATGGGTCACAAAGAGTGTGGACAC 2767
 Db 73163 TGGTGTATGAGACAGGAGGCGCTGCGATGCTGGGATTCATGGGTCGCAAGGAGTGTGGACAT 73222
 QY 2768 GACTCAGCAACTGAACTCAACTGAACTGTAAGAACCTTGTAGTATTTATATTACTCAGA 2827
 Db 73223 GACTCAGGAGCTGATCTGATCTGATCTGACTGACTACTACTTTCAGACTGTGTGCTGCTG 73282
 QY 2828 AAATAGTAAATTCATATATTCATAAATTTTTCATAATGTGTGTTAGATAAATAAGATT 2887
 Db 73283 TGAGGGCTACTCCATTTCTTCAAGGCACTTCTTGCACACAGTAGTAGATATAATAGTGCAT 73342
 QY 2888 TTCAAAATTTGATTTTATCTTTTCTCTACTTATTTTATTTTGGGATTTTAA 2943
 Db 73343 CTGATTTAAATTTTCCCATTCAGTCTATTTTGTCTGACTGATTTCCCAAAATTTGA 73398

RESULT 18

AC146889 139617 bp DNA linear HTG 04-NOV-2003
 Bos taurus clone RP42-509J20, WORKING DRAFT SEQUENCE, 4 ordered
 pieces.

AC146889
 AC146889 GI:38153770

HTG: HTGS_PHASE2; HTGS_DRAFT.

KEYWORDS
 Bos taurus (cow)

SOURCE
 Bos taurus

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

REFERENCE
 1 (bases 1 to 139617)

AUTHORS

Antonellis, A., Avele, K., Benjamin, B., Blakesley, R. W.,
 Bouffard, G. G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,
 Coleman, H., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurie, B., Idol, J. R.,
 Karlins, E., Kwong, P., Laric, P., Larson, S., Lee-Lin, S.-Q.,
 Legaspi, R., McDowell, J., Maduro, V. B., Margulies, E. H., Masiello, C.,
 Maskeri, B., McDowell, J., Mullikin, J. C., Paquirigan, C., Pearson, R.,
 Portnoy, M. E., Prasad, A., Reddix-Dugue, N., Schandler, K.,
 Schueler, M. G., Shah, K., Sison, C., Stantripop, S., Thomas, J. W.,
 Thomas, P. J., Tsipouri, V., Vogt, J. L., Wetherby, K. D., Young, A. and
 Green, E. D.

NISC Comparative Sequencing Initiative

Unpublished

REFERENCE
 2 (bases 1 to 139617)

AUTHORS

Green, E. D.
 Direct Submission
 Submitted (16-OCT-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA

REFERENCE
 3 (bases 1 to 139617)

AUTHORS

Green, E. D.
 Direct Submission
 Submitted (04-NOV-2003) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Nov 4, 2003 this sequence version replaced gi:37693553.

COMMENT

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc_zoo@hgr.nih.gov

----- Project Information

Center project name: eyr

Center clone name: 509J20

The sequence data in this record represents an 'enhanced' version of a phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g.,

human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 138458 bases at least Q40
 Consensus quality: 138035 bases at least Q30
 Consensus quality: 139121 bases at least Q20
 Insert size: 125600; agarose-fp
 Insert size: 139317; sum-of-contigs
 Quality coverage: 15.38x in Q20 bases; agarose-fp
 Quality coverage: 13.80x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 59777: contig of 59777 bp in length

* 59778: gap of unknown length

* 59778 114249: contig of 54372 bp in length

* 114250 114349: gap of unknown length

* 114350 135352: contig of 21003 bp in length

* 135353 135452: gap of unknown length

* 135453 139617: contig of 4165 bp in length.

FEATURES

source	1. 139617
	/organism="Bos taurus"
	/mol_type="genomic DNA"
	/db_xref="taxon:9913"
	/clone="RP42-509J20"
	/clone_lib="RP42"
misc_feature	1. 59777
	/note="assembly_fragment"
	clone_end:SP6
	vector_side:left
misc_feature	59778..114249
	/note="assembly_fragment"
misc_feature	114350..135352
	/note="assembly_fragment"
misc_feature	135453..139617
	/note="assembly_fragment"
	clone_end:T7
	vector_side:right
Query Match	58.08; Score 2071.6; DB 1; Length 139617;
Best Local Similarity	76.63; Pred.No. 0;
Matches 2596;	Conservative 0; Mismatches 169; Indels 626; Gaps 11;
Qy	1 GTCAGGAAGGCGGCGAGTGGAGAGGAGTACCTACCTCGTCCCAAGGTAAAGGACAGTAGC 60
Db	29610 GTCAGGAAGGTCGCGGTAGGAGA-----TACCCCTCGTCCAAAGGTAAAGGACAGTGGC 29664
Qy	61 TGGCCTTTGTGGAGCGCGGTAAAGAGATACCCAGCGCCCAAGGTAAAGAAACCCCAAG 120
Db	29665 TGTGCTTTGTGCGAGCGCGGTGAAGAGATACCCATCGCCCAAGGTAAAGAAACCCCAAG 29724
Qy	121 TAAGATGGTAGGTGTTCTGGAGAGGCGCATCAGAGGCGGAGACATCTGAAACCATACAGCA 180
Db	29725 TAAGATGGTAGGTGTTCTGGAGAGGCGCATCAGAGGCGGAGACATCTGAAACCATACAGCA 29784
Qy	181 GAAAACTAGTCAATCTAATCAGTACCTAGGACCAAGCGCTTGTCTAATCTCAATGAACTAAG 240
Db	29785 GAAAACTAGTCAATCTAATCAGTACCTAGGACCAAGCGCTTGTCTAATCTCAATGAACTAAG 29844

Qy	241 CCATGCCCGTGGGCAACCCCAAGATGGCAGGTATGTTGGAGAGATCTGACAGATGTG 300
Db	29845 CCATGCCCGTGGGCAACCCCAAGATGGCAGGTATGTTGGAGAGATCTGACAGATGTG 29904
Qy	301 GTCCACTGGAGAGGGAAT-GCAAAACCACTTCAGTATTCTTCCCTTGAGAAACCCCATGAA 359
Db	29905 GTCCACTGGAGAGGGAATGCAAAACCACTTCAGTATTCTTCCCTTGAGAAACCCCATGAA 29964
Qy	360 CAGTATGAAAAGGCAAAATAGTAGTACTGAAAGAGAACTCCCCAGGTCAAGTAGTGC 419
Db	29965 CAGTATGAAAAGGCAAAATAGTAGTACTGAAAGAGAACTCCCCAGGTCAAGTAGTGC 30021
Qy	420 CCCATATCTACTGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAGCC 479
Db	30022 -----TGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAGCC 30075
Qy	480 AAAGCAAAAGAATAACCCAGCTGGATGTGCTGCTGATATAAGCAAGGTCGAGTGTG 539
Db	30076 AGGGCAAAACAATACCCAGCTGGATGTGCTGCTGATATAAGCAAGGTCGAGTGTG 30135
Qy	540 TAAAGAGCAATATTGCATAGGAACCTGGAATGTCAGGTCCATGATCAAGCAAAATTGGA 599
Db	30136 TAAAGAGCAATATTGCATAGGAACCTGGAATGTCAGGTCCATGATCAAGCAAAATTGGA 30195
Qy	600 AGTGTCTCAACAAAGAGATGGCAAGATGCAATGTCAACATTTCTAGGAATCAGCGAACTAAA 659
Db	30196 AGTGTCTCAACAAAGAGATGGCAAGATGCAATGTCAACATTTCTAGGAATCAGCGAACTAAA 30255
Qy	660 ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTCGGGCAGGA 719
Db	30256 ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTACT-TTTTCTGGA 30313
Qy	720 ATCCCTCAGAGAAATGGAGTAGCCATCATGTCACAAAAGAGTCCGAAATGCAAGTACT 779
Db	30314 ATCCCTCAGAGAAATGGAGTAGCCATCATGTCACAAAAGAGTCCGAAATGCAAGTACT 30373
Qy	780 TGGATGCACTCTCAAAAACGACAGAAATCATCTGTGTTTGTTCCTCAAGGCAAAACCAATTCAA 839
Db	30374 TGGATGCACTCTCAAAAACGACAGAAATCATCTGTGTTTGTTCCTCAAGGCAAAACCAATTCAA 30433
Qy	840 TATCAGATTAATCCAGTCTATGCCCCCAACCAAGTAATGCTGAAGAGCTGAAGTTGAAG 899
Db	30434 TATCAGATTAATCCAGTCTATGCCCCCAACCAAGTAATGCTGAAGAGCTGAAGTTGAAG 30493
Qy	900 GTCCCTATGAAGACCTTACAGACCTTTTAAAGACTAACACCCCAAAAAGATGCTCTCTCAT 959
Db	30494 GTTCCTATGAACATCTACAGACCTTTTAAAGACTAACACCCCAAAAAGATGCTCTCTCTCAT 30553
Qy	960 TATAGGGACCTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGTAAACAGGCAAAAT 1019
Db	30554 TATAGGGACCTGGAATG-AAAAGTAGGATGTCAAGAAACACCTGGAGTAAACAGGCAAAAT 30612
Qy	1020 TGGCCTTGGAAATACGGAATGAAGCAGGGCAAGACTAATAGAGTTTTCCTCAAGAAATATGC 1079
Db	30613 TGGCCTTGGAAATGAAGTAGGAAGCAAGGCAAGACTAATAGAGTTTTCCTCAAGAAATATGC 30672
Qy	1080 ACTGTCTATAGCAAAACACCTCTTCCAAACAACAAGAGAGAGCTCTACATGAGACATC 1139
Db	30673 ACTGTCTATAGCAAAACACCTCTTCCAAACAACAAGAGAGAGCTCTACATGAGACATC 30732
Qy	1140 ACCAGATGGTCAACACCGAAATCAGATTGATTATATCTTTTGCAGCCCAAGAGTGGAGAG 1199
Db	30733 ACCAGATGGTCAACACTTAAATCAGATTGATTATATCTTTTGCAGCCCAAGAGTGGAGAG 30792
Qy	1200 CTCTATACAGTCAGCAAAAACCAAGACCCAGGAGCTTACTGTGGCTCAGATCATGAACTCCT 1259
Db	30793 CTCTATACAGTCACAAAACCAAGACCCAGGAGCTGACTGTGGCTCAGATCATGAACTCCT 30852
Qy	1260 TATTGCCAAAATTCAGACTTAAATGAAGAAAGTAGGGGAAACCACTAGATCACTCAGGTA 1319
Db	30853 TATTGCCAAAATTCAGACTTAAATGAAGAAAGTAGGGGAAACCACTAGATCACTCAGGTA 30912

Qy	1320	AGACCTAAATCCAAATCCCTTATGATATATACAGTGGAAATGAGAAATAGATTTTAAGGGCCT	13179
Db	30913	TGACCTAAATCAAAATCCCTTATGATATATACAGTGGAAATGAGAAATAGATTTTAAGGGTCT	30972
Qy	1380	AGATCTGATAGACAGAGTACCTAATCAACTATGGACAGAGGTTTCATGACATTTGTACAGGA	1439
Db	30973	AGATCTGATAGATAGAGTGCCCTGATGAACTATGGAAATGAGGTTCTGACATTTGTATAGGA	31032
Qy	1440	GACAGGATCGAGACCAATCCCATGGAAGAAATGCAAAAAAGCAAAATGGCTGTCCTGG	1499
Db	31033	GACAGGATCAAGACCAATCCCATGGAAGAAATGCAAAAAAGTAAATGGCTGTCCTGG	31092
Qy	1500	GGAGGCTTACAAATAGCTCTGAAAGAGAGAGAGTGAAGAAACCAAGCAAGAAAAAGGAAAG	1559
Db	31093	GGAGGCTTACAATAGTTGTGAAAAGAGAGAG-CAAAAAACAAGAGAAAAAGAAAT	31151
Qy	1560	ATAAAGCATCTGAATGCGAGGTTCCAAA-	1588
Db	31152	ATAAAGCATCTGAATGTAGAGTTTCCAAAAGATAGCAAGAGAGATAAGAAAGCCTTCCT	31211
Qy	1589	-----	1588
Db	31212	CAGTGATCAATCAAAAGAAATAGAGAAAAATAGAGAAATGGGAAAGACTAGAGATCTCTT	31271
Qy	1589	-----	1588
Db	31272	CAAGAAATCAGAGATACCAGGGAAACATTTTCATGCAAGATGGCTCGATAAAGGACAG	31331
Qy	1589	-----	1588
Db	31332	AAATGGTATGGACCTTAACAGAAGCAGAGATATTAAGAGAGGTGGCAAGAAATACACAGA	31391
Qy	1589	-----	1588
Db	31392	AGAACTGCACAAAAAGAGCTTCACGACACAGATATCACGATGGTGTGATCATCTGACCT	31451
Qy	1589	-----	1588
Db	31452	AGAGCCAGACATCCTGGAATGTGAAGTCAAGTCAAGTGGGCTTAGAAAAGCATCACTAGAACAA	31511
Qy	1589	-----	1588
Db	31512	AGTAGTGGAGTGATGGAAATCCAGCTTGAGCTATTTCAATCTCTGAAAGATGATGCTGT	31571
Qy	1589	-----	1588
Db	31572	GAAAGTCTGCACTCAATATGCCAGCAAAATTTGGAAAATCAGCGTGGCCACAGGACTG	31631
Qy	1589	-----	1588
Db	31632	GAAAGGTCAGTTTTCATTCCAATCCCAAGAAGGCAATGCCAAGAATGCTCAAACTA	31691
Qy	1589	-----	1588
Db	31692	CTGCACAAATTGCACATCTCACACACTAGTAAAGTAATGCTCAAAATGCTCCAAGCCAG	31751
Qy	1589	-----GAACTCCAGTTGTTCAAGCTGGTTTTAGAAAAGTCAAG	1626
Db	31752	GCTTTAGCAATACGTGAACCGTGAACTCCAGATGTTCAAGCTGGTTTTAGAAAAGGCGAG	31811
Qy	1627	AGAAACAGAGACCAATTTGCCAATCTCTGTATCATGGAAGAAGCAAGAGAGTTCCA	1686
Db	31812	AGGAACGAGATCAAAATTTGCCAATCCACCTGGATCATGGAAGAAGCAAGAGAGTTCCA	31871
Qy	1687	GAAAAACATCTATTTCTGCTTTATTGACTATGCAAAAGCCTTTGACTGTGGGGTCAAA	1746
Db	31872	GGAAAAACATCTATTTCTGCTTTATTGACTATGCCAAAGCCTTTGACTGTGTGATCAAA	31931
Qy	1747	TAAACTGTGNAATTTCTCAAAAGGATGGGAATCCAGACCCTGACTGACTGACTTGAA	1806
Db	31932	TAAACTGTGNAATTTCTCAAAAGATGGGAATCCAGACCCTGACTGACTGCTCTTTGAG	31991
Qy	1807	AAATTTGTATGAGGTCAGAAAGCAACAGTTAGAACTGGAATGGAACCAACAGACTGGTT	1866

[illegible]

RESULT 19
AC145440/c
LOCUS
DEFINITION

Db	147223	CAAAATGGAAGTGGTCAAAACAGAGATGGCAAGAGTGAATGTTGACATTTCTAGGAATCGG	147164
QY	651	CGAACTAAATGGACTGGATGGGTGAATTTAACTCAGATGACCATTTATCTACTACTG	710
Db	147163	CAAACTAAATGGACTGGATGGGTGAATTTAACTCAGATGACCATTTATCTACTACTG	147104
QY	711	CGGGCAGGAATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCCGAAA	770
Db	147103	TGGCAGGAATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCTCAAG	147044
QY	771	TGCAGTACTTGGATGCGAGTCTCAAAAAGCAGAGAAATGATCTCTGTTTGGTTCCAGGCCAA	830
Db	147043	TGCAGTACTTGGATGCGAGTCTCAAAAAGCAGAGAAATGATCTCTGTTTGGTTCCAGGCCAA	146984
QY	831	ACCAATTCATATACAGTAATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAGAAGCTGA	890
Db	146983	ACCAATTCATATACAGTAATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAGAAGCTGA	146924
QY	891	AGTTGAAGGCTCTATGAAGACCTTACAGACCTTTTGAAGCTTAACACCCCAAAAAGATGT	950
Db	146923	AGTTGAAGGCTCTATGAAGACCTTACAGACCTTTTGAAGCTTAACACCCCAAAAAGATGT	146864
QY	951	CTTCTCATATAGGGGACTGGATGCAAAAGTAGGAAGCAAGACACCTGGAGTAAC	1010
Db	146863	CTTCTCATATAGGGGACTGGATGCAAAAGTAGGAAGCAAGACACCTGGAGTAAC	146804
QY	1011	AGGCAAAATTTGGCCTTGGATACGGAAATGAAGCAGGGCAAGACTAATAGAGTTTGGCA	1070
Db	146803	AGGCAAAATTTGGCCTTGGATACGGAAATGAAGCAGGGCAAGACTAATAGAGTTTGGCA	146744
QY	1071	AGAAATGCACTGCTATAGCAACACCTCTTCCACACACACAGAGAGAGCTCTACAC	1130
Db	146743	AGAAATGCACTGCTATAGCAACACCTCTTCCACACACACAGAGAGAGCTCTACAC	146684
QY	1131	ATGGACATCACAGATGGTCAACACCGAAATCAGATTGATTATATCTTTGCCAGCCAAAG	1190
Db	146683	ATGGACATCACAGATGGTCAACACCGAAATCAGATTGATTATATCTTTGCCAGCCAAAG	146624
QY	1191	ATGGAGAGGCTCTATACAGTCAGCAAAAACAGACCGAGGCTTACTGTGGCTCAGATCA	1250
Db	146623	ATGGAGAGGCTCTATACAGTCAGCAAAAACAGACCGAGGCTTACTGTGGCTCAGATCA	146564
QY	1251	TGAATCTCTTATTCGCAAAATTCAGACTTAAATTCAGCAAAAGTAGGGGAAAACCACTAGATC	1310
Db	146563	TGAATCTCTTATTCGCAAAATTCAGACTTAAATTCAGCAAAAGTAGGGGAAAACCACTAGATC	146504
QY	1311	ACTCAGGTAAAGCTTAAATCCATCCCTTATGATTATACAGTGGAAAGTAGAGAAATAGATT	1370
Db	146503	ACTCAGGTAAAGCTTAAATCCATCCCTTATGATTATACAGTGGAAAGTAGAGAAATAGATT	146444
QY	1371	TAAAGGCTAGATCTGATAGACAGTACCTTAATGAACTATGGACAGAGCTTATGACAT	1430
Db	146443	TAAAGGCTAGATCTGATAGACAGTACCTTAATGAACTATGGACAGAGCTTATGACAT	146384
QY	1431	TGTACAGAGACAGGGATCGAGACCATCCCATCGAAAAGAAATGCAAAAAGCAAAATG	1490
Db	146383	TGTACAGAGACAGGGATCGAGACCATCCCATCGAAAAGAAATGCAAAAAGCAAAATG	146324
QY	1491	GCTGTCTGGGAGGCTTAAATAGCTGTGAAGAAAGAGAGTGAAGACCAAGGAAA	1550
Db	146323	GCTGTCTGGGAGGCTTAAATAGCTGTGAAGAAAGAGAGTGAAGACCAAGGAAA	146264
QY	1551	AAAGGAAGATAAAAGCATCTGAATCGAGAGTCCAAA	1588
Db	146263	AAAGGAAGATAAAAGCATCTGAATCGAGAGTCCAAA	146204
QY	1589	-----	1588
Db	146203	AGCCTTCTCAGTGATCAATGCAAAAGAAATAGAGGAAAACACACAGAAATGGGAAAGACTAG	146144
QY	1589	-----	1588
Db	146143	AGATCTCTCAAGAAATTAGAGATACCAAGGGACATTTTCATGCAAAAGATGGCTCGAT	146084
QY	1589	-----	1588
Db	146083	AAAGCAGAAATGATATGGACCTTAACAGACGACAGATATTAAGAAGAGATGGCRAAGA	146024
QY	1589	-----	1588
Db	146023	ATACACAGAAAGAACTATACAAAAAGATCTTTCATGACCCAGATATACAGATGATGAT	145964
QY	1589	-----	1588
Db	145963	CACTGACCTAGAGCCAGACATCCTCGAATGTGAATCAAGTGGGCTTTAGAAAGCATCAC	145904
QY	1589	-----	1588
Db	145903	TATGAACAAAGCCAGTGGAGGTGATGGAATTCAGTTGAGCTCTTTCAATACTGAAAGA	145844
QY	1589	-----	1588
Db	145843	TGATGCTGTGAAGTGTTCGACTCAATATGCCAGCAAAATTTGGAACACACAGCAGTGCC	145784
QY	1589	-----	1588
Db	145783	ACAGGCTGGAAGAGGTGAGTTTTCATTCCAATCCCAAGAAAGCAANTGCCAAAGATG	145724
QY	1589	-----	1588
Db	145723	CTCAAACTACCGCACAAATTTGCACATCATCTCGCATGCTAGTAAAGTAATGCTCAAAATCT	145664
QY	1589	-----	1617
Db	145663	CCAGGCCAGGCTTCAGCAATACGTGAACCGTGAATTTCTGATGTTCAAGCTGGTTTTAG	145604
QY	1618	AAAACTCAGAGAAACACAGAGCCAAATTTGCCAACATCTCTGTATCATATGAAAAAGCAAG	1677
Db	145603	AAAACTCAGAGAAACACAGAGCCAAATTTGCCAACATCTCTGTATCATATGAAAAAGCAAG	145544
QY	1678	AGAGTTCCAGAAAAACATCTATTTCTGCTTTTATGACTATCAAAAGCCTTTGACTGTGG	1737
Db	145543	AGAGTTCCAGAAAAACATCTATTTCTGCTTTTATGACTATCAAAAGCCTTTGACTGTGG	145484
QY	1738	GGGTCAATAACTGTGAAAAATTTCTGAAAGGGATGGGAATACCAAGACCACTGACCTG	1797
Db	145483	GGGTCAATAACTGTGAAAAATTTCTGAAAGGGATGGGAATACCAAGACCACTGACCTG	145424
QY	1798	ACTTTGAAAAATTTGTATGAGGTGAGAGCAACAGTTAGAACTGACATGGAACCAAC	1857
Db	145423	ACTTTGAAAAATTTGTATGAGGTGAGAGCAACAGTTAGAACTGACATGGAACCAAC	145364
QY	1858	AGACTGGTTTCAAGTAGGAAAAAGGAGTATGTCAAGGCTGTATATTGTCAACCCGGCTTGT	1917
Db	145363	AGACTGGTTTCAAGTAGGAAAAAGGAGTATGTCAAGGCTGTATATTGTCAACCCGGCTTGT	145304
QY	1918	TAACTTTATGAGAG-A-CATCATGAGAAACGCTGGGTGGAAGCAAGCAAGCTGGAAT	1976
Db	145303	TAACTTTATGAGAGTACATCATGAGAAACACACTGGGTGGAAGCAAGCAAGCTGGAAT	145244
QY	1977	CAAGATTCCGGGAGAAATAGCAATTAACCTCAGATATGAGATGATACCACTTATGGC	2036
Db	145243	CAAGATTCCGGGAGAAATAGCAATTAACCTCAGATATGAGATGATACCACTTATGGC	145184
QY	2037	AGAAATGAAGAGAACTAAAAAGCCTTCTGATGAAGGTGAAAGGAGAGATGAAAAAGT	2096
Db	145183	AGAAATGAAGAGAACTAAAAAGCCTTCTGATGAAGGTGAAAGGAGAGATGAAAAAGT	145124
QY	2097	TGGCTTAAAGCTCAACATTCAGAAACCAAGATCATGTCATCTGTCCTCCATCATCTCATG	2156
Db	145123	TGGCTTAAAGCTCAACATTCAGAAACCAAGATCATGTCATCTGTCCTCCATCATCTCATG	145064
QY	2157	GGAAATAGATGGGAAAAACAGTGGAAACAGTGTGAGACTTTATTTTGGGGGGCTCCAAA	2216
Db	145063	GGAAATAGATGGGAAAAACAGTGGAAACAGTGTGAGACTTTATTTTGGGGCTCCAAA	145015

QY 2217 ATCACTGACATGGTACTGCACCCATGAATTAATAAGACACATTACTCTTGGAGAAAA 2276
Db 145014 CTCACCTGACATGGTACTGCACCCATGAATTAATAAGACGCTTACTCTTGGAGAAAA 144955
QY 2277 GTTA--ACCACTAGATACATATGAAAAGCAGACATTAACCTTGGCCACAAAGCC 2334
Db 144954 GTTATGACCACTAGATGATATTAATAAGCAGACATTAACCTTGGCCACAAATGTC 144895
QY 2335 CATCTAGTCAAGCTATGGTGTCTTCCAGTGGTCTATGATGATGATGATGATGATGATG 2394
Db 144894 CTCTAGTCAAGCTATGGTGTCTTCCAGTGGTCTATGATGATGATGATGATGATGATG 144835
QY 2395 AGAAAGCTGAGCAGCAGAAATGATGCTTTTGAAGTGGTGGTGGAGAACTCTTG 2454
Db 144834 AGAAAGCTGAGCAGCAGAAATGATGCTTTTGAAGTGGTGGTGGAGAACTCTTG 144775
QY 2455 AGAGTCCCTTGGAGCTCAAGAGATCAACAGTCCATCTGAAGAGATCAGCCCTGGG 2514
Db 144774 AGAGTCCCTTGGAGCTCAAGAGATCAACAGTCCATCTGAAGAGATCAGCCCTGGG 144715
QY 2515 ATTCTTTTGAAGAAATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGA 2574
Db 144714 ATTCTTTTGAAGAAATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGA 144655
QY 2575 AGAGTCACTCACTGGAAGAGCCCTGATGCTGGAGGATGGGGCAGGAGAGAGG 2634
Db 144654 AGAGTCACTCACTGGAAGAGCCCTGATGCTGGAGGATGGGGCAGGAGAGAGG 144595
QY 2635 GGACGACAGAGGATGAGATGGCTGGATGGGATCAGTCACTGATGATGATGATGATGATG 2694
Db 144594 GGACGACAGAGGATGAGATGGCTGGATGGGATCAGTCACTGATGATGATGATGATG 144535
QY 2695 GAATCTCTGAGTGGTGTGAGCAGGAGGCTCTCTCGCGGATCTATGGGGTACAA 2754
Db 144534 GAATCTCTGAGTGGTGTGAGCAGGAGGCTCTCTCGCGGATCTATGGGGTACAA 144475
QY 2755 AAGAGTGGACACACTGAGCACTGAACTGAACTGAACT 2794
Db 144474 AAGAGTGGACACACTGAGCACTGAACTGAACTGAACTGAACTGAACT 144435

RESULT 20
AC144498/c
LOCUS AC144498 294688 bp DNA linear HTG 02-OCT-2003
DEFINITION Bos taurus clone RP42-25714, WORKING DRAFT SEQUENCE, 9 ordered
pieces.
ACCESSION AC144498
VERSION AC144498.3 GI:37360989
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 294688)
AUTHORS Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K.,
Chen, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X.,
Gupta, J., Haghigat, P., Han, J., Hansen, N., Ho, S.-L., Hu, P.,
Hurler, B., Idol, J.R., Karlins, E., Kwong, P., Latic, P., Lee, J., Lee, S.-Q.,
Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Mastaglio, C.,
Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E.,
Prasad, A., Reddix-Dugue, N., Schardler, K., Schueler, M.G., Shah, K.,
Sison, C., Stantirip, S., Thomas, J.W., Thomas, P.J., Tsipouri, V.,
Vogt, J.L., Wetherby, K.D., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
REFERENCE 2 (bases 1 to 294688)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2003) NIH Intramural Sequencing Center, 8717
Grosvont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 294688)

AUTHORS
TITLE
JOURNAL
COMMENT

Green, E.D.
Direct Submission
Submitted (02-OCT-2003) NIH Intramural Sequencing Center, 8717
Grosvont Circle, Gaithersburg, MD 20877, USA
On Oct 2, 2003 this sequence version replaced gi:31621329.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hgr.nih.gov
----- Project Information
Center project name: eem
Center clone name: 257104

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 293115 bases at least Q40
Consensus quality: 293600 bases at least Q30
Consensus quality: 293797 bases at least Q20
Insert size: 187000; agarose-fp
Insert size: 293888; sum-of-contigs
Quality coverage: 17.56x in Q20 bases; agarose-fp
Quality coverage: 11.18x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.

1 1193 1192: contig of 1192 bp in length
1293 1292: gap of unknown length
52164 52164: contig of 50872 bp in length
52165 52164: gap of unknown length
52265 71131: contig of 18867 bp in length
71132 71131: gap of unknown length
71232 77081: contig of 5850 bp in length
77082 77181: gap of unknown length
77182 81753: contig of 4572 bp in length
81754 81853: gap of unknown length
81854 106709: contig of 24856 bp in length
106710 106809: gap of unknown length
130077 130176: gap of unknown length
130177 282832: contig of 152656 bp in length
282833 282932: gap of unknown length
282933 294688: contig of 11756 bp in length.

FEATURES
source

1..294688
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-25714"
/clone_lib="RP42"
1..1192
/note="assembly_fragment
misc_feature

```

clone_end:SP6
vector_side:left
misc_feature 1293..52164
  /note="assembly_fragment"
misc_feature 52265..71131
  /note="assembly_fragment"
misc_feature 71232..77081
  /note="assembly_fragment"
misc_feature 77182..81753
  /note="assembly_fragment"
misc_feature 81854..106709
  /note="assembly_fragment"
misc_feature 106810..1130076
  /note="assembly_fragment"
misc_feature 130177..282832
  /note="assembly_fragment"
misc_feature 239910..294688
  /note="clone overlaps with GenBank Accession Number
  AC144602 clone RP42-212N23 (center project name een)"
misc_feature 282933..294688
  /note="assembly_fragment
clone_end:right
vector_side:right

Query Match      57.8%; Score 2066.1; DB 1; Length 294688;
Best Local Similarity 76.5%; Pred. No. 0;
Matches 2544; Conservative 0; Mismatches 179; Indels 601; Gaps 7;

QY 126 TGGTAGGTTGTGAGAGGGGATCAGAGGGGAGACNCTACTGAAACCATACACGGAGAAA 185
DB 85445 TCGCAGGTTGTCGAGGAGGGGATCAGAGGGGAAAACACACTGAAACCATACTCAGAAA 85386
QY 186 CTAGTCAATCTAATCACACTAGGACCACAGCTTGTCTAATCAATGAACTAAGCCATG 245
DB 85385 CTAGTCAATCTAATCACACTAGGACCACAGCTTGTCTAATCAATGAACTAAGCCATG 85326
QY 246 CCCTGGGGCCAAACCAAGATGGGAGGTCATGTTGGAGAGATCTGACAGAAATGTGTCCA 305
DB 85325 CCAGTGGGGCCAAACCAAGATGGGCGGTACGGTGGAAAGATCTGACAGAAATGTGTCCA 85266
QY 306 CTGGAGAGGGAAT-GCAACCACTTCAGTATCTTCCTCTGAAACCCCATGAACAGTA 364
DB 85265 CTGGAGAGGGAATGGCAACCACTTCAGTATCTTCCTCTGAAACCCCATGAACAGTA 85206
QY 365 TGAAGAGCCAAATCATAGGATCTGAAAGAGGAACTCCCCAGGTCACTAGTGGTCCCAT 424
DB 85205 TGAAGAGCCAAATCATAGGATCTGAAAGAGGAACTCCCCAGGTCACTAGTGGTCCCAT 85146
QY 425 ATGCTACTGGAGATCAGTGGAGAAATTAATCTCCAGAAAGATGAAGATCGAGCCAAAGC 484
DB 85145 ATGCTACTGGAGATCAGTGGAGAAATTAATCTCCAGAAAGATGAAGATCGAGCCAAAGC 85086
QY 485 AAAAAGAAATACCAGCTGTGATGTGACTGTGTATATAAGCAAGTCCGATGCTGTAAAG 544
DB 85085 AAAAAGAAATACCAGCTGTGATGTGACTGTGTATATAAGCAAGTCCGATGCTGTAAAG 85026
QY 545 AGCAATATTCATAGGAACCTTGAATGTCTAGGTCCATGAATCAAGGCAAAATGGAAGTGG 604
DB 85025 AGCAATATTCATAGGAACCTTGAATGTCTAGGTCCATGAATCAAGGCAAAATGGAAGTGG 84966
QY 605 TCAACCAAGATCGCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAATCTAAATGGA 664
DB 84965 TTAAGCAAGATCGCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAATCTAAATGGA 84906
QY 665 CTGAAATGGGTGAATTTAACTCAGATGAACCAATATATCTACTCTGCGGACGAATCCC 724
DB 84905 CTGAAATGGGTGAATTTAACTCAGATGAACCAATATATCTACTCTGCGGACGAATCCC 84846
QY 725 TCAGAGAAATGGAGTACCATCTGCTCAACAAAGAGTCCGAAATGCAAGTACTTGGAT 784
DB 84845 TCAGAGAAATGGAGTACCATCTGCTCAACAAAGAGTCCGAAATGCAAGTACTTGGAT 84786
QY 785 GCAGTCTCAAAACCGACAGAAATGATCTCTGTTGTTGTTTCCAAAGGCAAACTTCAATATCA 844

```

```

DB 84785 GCAATCTCAAAATGACAGATGATCTCTGTTCAATTTCAAGGCAAACTTCAATATCA 84726
QY 845 CAGTAATCCAAGTCTATGCCCAACCAAGTAATGTCTGAAAGAGCTGAAGTGAACGGTCT 904
DB 84725 CAGTAATCCAAGTCTATGCCCAACCAAGTAATGTCTGAAAGAGCTGAAGTGAACGGTCT 84666
QY 905 ATGAAGACCTACAGACCTTTTAGAATTAACACCCCAAAAGAGATGCTCTTCTCATTATAG 964
DB 84665 ATGAAGACCTACAGACCTTTTAGAATTAACACCCCAAAAGAGATGCTCTTCTCATTATAG 84606
QY 965 GGGACTTGGAAATGAAAGTAGGAAGCAAAAGCAACCTGGAGTAACAGGCAAAATTTGGCC 1024
DB 84605 GGGACTTGGAAATGAAAGTAGGAAGCAAAAGCAACCTGGAGTAACAGGCAAAATTTGGCC 84546
QY 1025 TTGGAATACGGAATGAAGCAGGGCAAAAGACTTAATAGAGTTTGGCAAGAAATGCACTGG 1084
DB 84545 TTGGAATACGGAATGAAGCAGGGTAAGAGCTAATAGAGTTTGGCAAGAAATGCACTGG 84486
QY 1085 TCATAGCAAAACCCCTCTTCCAAACCAACAGAGAGAGACTCTACACATGGACATCACCAG 1144
DB 84485 TCATAGCAAAACCCCTCTTCCAAACCAACAGAGAGAGACTCTACACATGGACATCACCAG 84426
QY 1145 ATGGTCAACACCGAAATCAGATTGATTATATTTTTCAGGCCAAAGATGGAGAACTCTTA 1204
DB 84425 ATGGTCAACATCGAAATCAGATTGATTATATTTTTCAGGCCAAAGATGGAGAACTCTTA 84366
QY 1205 TACAGTCAGCAAAACCAAGCAGGCTTACTGTGGCTCAGATCATCACTCCTTATTG 1264
DB 84365 TACAGTCAGCAAAACCAAGCAGGCTTACTGTGGCTCAGATCATCACTCCTTATTG 84306
QY 1265 CCAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAAAGACC 1324
DB 84305 CCAATTCAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAAAGACC 84246
QY 1325 TAAATCCAAATCCCTTATGATTATACAGTGGAGATGAGAAATAGATTTAAGGGCCTAGATC 1384
DB 84245 TAAATCCAAATCCCTTATGATTATACAGTGGAAATGAGAAATAGATTTAAGGGCCTAGATC 84186
QY 1385 TCATAGACAGAGTACTTAAATGATGATGACAGAGGTTTCATGACATTTTACAGAGACAG 1444
DB 84185 TCATAGATAGAGTGTCTGATGACTTATGGAATGAGTTTCATGACATTTTACAGAGACAG 84126
QY 1445 GGATCGAGACCATCCCATGGAAAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGGAGG 1504
DB 84125 GGATCAAGACCATCCCATGGAAAAGAAAGCAAAAAGCAAAATGGCTGTCTGGGGAGG 84066
QY 1505 CCTTACAAATAGCTGTGAAAAGAGAGAGTCAAAAGCAAAAGGAAAAGGAAAAGATATAA 1564
DB 84065 CCTTACAAATAGCTGTGAAAAGAGAGAGTCAAAAGCAAAAGGAAAAGGAAAAGATATA 84006
QY 1565 AGCATCTGAATGACAGGTTTCCAAAG- 1589
DB 84005 AACATCTGAATGACAGGTTTCCAAAGAGATTAAGAGAGATTAAGAAAGCCTTCTTCACTG 83946
QY 1590 ----- 1589
DB 83945 ATCAATGCAAGAAATAGAGGAAAACAAAGATGGGAAAGACTAGGAATCTCTTCAAGA 83886
QY 1590 ----- 1589
DB 83885 AAATCAGAGATACCAAGGAACATTTTCATGCAAGATGAGCTCGATAAAGGACAGAAATG 83826
QY 1590 ----- 1589
DB 83825 GTATGGATCTTAACAGAGCAGAGATATTAAGAGAGATGGCAAGAAATACACAGAGAAC 83766
QY 1590 ----- 1589
DB 83765 TGTAACAAAAGATCTTCCAGACCCAGATAATCAGATGGTGTGGTCACTGACCTAGAGC 83706
QY 1590 ----- 1589

```

Db 83705 TAGAAATCCTGGAAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTACAAACAAAGCTA 83646
Qy 1590 ----- 1589
Db 83645 GTGAGGTGATGGAATCCAACTGAGCTATTTCAAATCCTGAAAGATGATGCTGTGAAG 83586
Qy 1590 ----- 1589
Db 83585 TGCTGCACTCAATATGCCAGCAAAATTTGGAAACTCAGAGTGGCCACAGGATTTGGAATA 83526
Qy 1590 ----- 1589
Db 83525 GGTGAGTTTTCATTCCTCAATCGCGAAGAAAGCAATGCCAAGATGCTCAAACTACCGCA 83466
Qy 1590 ----- 1589
Db 83465 CAATTGCACTCATCTCACATGCTAGTAAAGTAATGCTCAAAATTTCTCAAGCCAGGCTTC 83406
Qy 1590 ----- 1631
Db 83405 AGCAATATGTGAACCGTCACTTCTGATGTTTCAAGCTGGTTTTAGAAAGGACAGAGAA 83346
Qy 1632 CCAGAGACCAAAATGCCAACATCTCTGTATCATATGGAAGCAAGAGATTTCCAGAAAA 1691
Db 83345 CCAGAGATCAAAATGCCAACATCTCTGGATCATGGAAGAAAGCAAGAGATTTCCAGAAAA 83286
Qy 1692 ACATCTATTTCTGCTTTATTTGACTATGCAGAGGCTTTGACTGTGGGGTTCACAAATTAAC 1751
Db 83285 ACATCTATTTCTGCTTTATTTGACTATGCAGAGGCTTTGACTGTGGGATCACAATTAAC 83226
Qy 1752 TGTGGAATTTCTGAAAGGATGGGAATACAGACCACTGACCTGACTCTTGAATAAT 1811
Db 83225 TGTGGAATTTCTGAAAGATAGGAATACAGACCACTGATCTGCCCCCTTGAAAT 83166
Qy 1812 TGTATGAGTTCAGAGACCAACAGTTAGACTGACATGGAACAAAGACTGGTTTCAAG 1871
Db 83165 TGTATGAGTTCAGAGACCAACAGTTAGACTGACATGGAACAAAGACTGGTTTCAAG 83106
Qy 1872 TAGAAAAAGGATGTGCAAGGCTGTATTTGTCACCCGCTGTTTAACTTCTATGSCAG 1931
Db 83105 TAGAAAAAGGATGTGCAAGGCTGTAGTTGTCACCCCTGTTTAACTTATTAAGCAG 83046
Qy 1932 AG-ACATCATGAGAAACGCTGGGTGGAAG-AAGCACAAGCTGGAATCAAGATTTGCGGG 1989
Db 83045 AGTACATCATGAGAAACGTTGGGCTGGAAGAAACCAACGCTGGAATCAAGATTTGCTGG 82986
Qy 1990 AGAATAGCANTACCTCAGATATGCAGATGATACCACTTATGGCAGAAAGTGAAG 2049
Db 82985 AGAATATCAATAACCTTAGATATGCAAGTACACCACTTATGGCAGAAAGTGAAG 82926
Qy 2050 GAACTAAAAAGCCTCTTGTATGAAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTC 2109
Db 82925 GAACTAAAAAGCCTCTTGTATGAAGTGAAGTGAAGTGAAGAAAGTTGGCTTAAAGCTC 82866
Qy 2110 AACATTCAGAAAAAGAGATCATGGCATCTGGTCCCATCTCATATGGGAAATAGATGGG 2169
Db 82865 AACATTCAGAAAAAGAGATCATGGCATCTGGTCCCATCTCATATGGGAAATAGATGGG 82806
Qy 2170 GAAACAGTGGAAACAGATGTGAGATTTATTTTGGGGGGCTCCAAATCACTGCAGATG 2229
Db 82805 GAAACAGTGGAAACAGATGTGAGATTTA--TTTGGGGGGCTCCAAATCACTGCAGATG 82748
Qy 2230 GTGACTGAGCCATGAAATTAAGACACTTACTCTTGGGAAAGAAAGTTA--ACCAACC 2287
Db 82747 GTGACTGAGCCATGAAATTAAGAGATGCTCTCTTGGGAAAGAAAGTTATGACCAACC 82688
Qy 2288 TAGATACATATTGAAAGAGAGACATTAACCTTGGCCAAAGAGCCCATCTAGTCAAGG 2347
Db 82687 TAGACAGCATATTCAAAAGAGAGACATTAACCTTGGCCAAAGAGATGCTCATCTAGTCAAG 82628
Qy 2348 CTATGGTTTTTCCAGTGGTCAATGATGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 2407
Db 82627 CTATGGTTTTTCCAGTGGTCAATGATGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 82569

Qy 2408 ACTGAGAAATGATGCTTTTGAACCTGTGTGTGGAGAGACTCTTTGAGAGTCCCTTGA 2467
Db 82567 GCCAAAGAAATGATGCTTTTGAACCTGTGTGTGGAGAGACTCTTTGAGAGTCCCTTGA 82508
Qy 2468 CTGCAAGAGATCCCAACAGTCCATCTCAAGAGATCAGCCCTGGGATTTCTTTGGAG 2527
Db 82507 CTGCAAGAGATCCCAACAGTCCATCTCAAGAGATCAGCCCTGGGATTTCTTTGGAG 82448
Qy 2528 GAATGATGCTAAAGCTGAAACCTCCAGTACTTTGGCCACCTGATCAGAAAGCTGACTCAC 2587
Db 82447 GAATGATGCTAAAGCTGAAACCTCCAGTACTTTGGCCACCTGATGAGAGTTGACTCAT 82388
Qy 2588 TGAAGAAGACCTGATGCTGGAGGATTTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAG 2647
Db 82387 TGAAGAAGACCTGATGCTGGAGGATTTGGGGGAGAGAGAGAGAGAGAGAGAGAGAG 82328
Qy 2648 TGAGATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2707
Db 82327 TGAGATGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 82268
Qy 2708 TGGTATGATGAGAGGAGGCTCTCTCTGGGGGCAATCATGGGGTCAAAAGATTTGGACAC 2767
Db 82267 TGGTATGATGAGAGGAGGCTCTCTCTGGGGTCTGCGATCATGGGGTGGCAAGAGTCA 82208
Qy 2768 GACTGAGCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 2827
Db 82207 AACTGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 82149
Qy 2828 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2851
Db 82148 GGTAAAGATGCTGCTTCAATGCA 82125

RESULT 21

AC135216
LOCUS AC135216
DEFINITION Bos taurus clone RP42-427020, WORKING DRAFT SEQUENCE, 3 ordered
pieces.

AC135216
VERSION AC135216.2 GI:25699761

KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

1 (bases 1 to 153264)

Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.I., Maduro,V.B.,
Margulies,E.H., Masello,C., Maskeri,B., McBowell,J.,
Paguigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Rendix-Dupre,N., Schandler,K., Schueler,M.G., Sison,C.,
Standridge,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 153264)

Green,E.D.

Direct Submission

Submitted (09-OCT-2002) NIH Intramural Sequencing Center, 8717

Grovermont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 153264)

Green,E.D.

Direct Submission

Submitted (27-NOV-2002) NIH Intramural Sequencing Center, 8717

Grovermont Circle, Gaithersburg, MD 20877, USA

On Nov 27, 2002 this sequence version replaced gi:23622199.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nhgri.nih.gov
----- Project information
Center project name: djz
Center clone name: 427020

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a phrap-derived quality score.

with a Phrag-derived quality score

----- Summary Statistics

----- Summary statistics -----

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.930319

Consensus quality: 152954 bases at least Q40

Consensus quality: 153013 bases at least Q30

Consensus quality: 153042 bases at least Q20

Insert size: 140000; agarose-ftp

Insert size: 153067; sum-of-contigs

Quality coverage: 11.86x in Q20 bases; agarose-ftp

Quality coverage: 10.94x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.

- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

```

* * the accession number will be preserved.
* 1 69854: contig of 69854 bp in length
* 69855 69954: gap of unknown length
* 69955 108944: contig of 38990 bp in length
* 108945 109044: gap of unknown length
* 109045 153264: contig of 44220 bp in length

```

FEATURES

```

FEATURES
  source
    1..153264
      Location/Qualifiers
        organism="Bos taurus"
        mol_type="genomic DNA"
        db_xref="taxon:9913"
        clone="RP42-427020"
        clone_lib="RP42"
      misc_feature
        1..69854
          /note="assembly_fragment"
          clone_end:T7
          vector_side:left
        misc_feature
          69955..108944
            /note="assembly_fragment"
            109045..153264
              /note="assembly_fragment"
              clone_end:SP6
              vector_side:right
    Query Match 57.8%; Score 2065.3; DB 1; Length 153264;
    Best Local Similarity 76.8%; Pred.No. 0;
    Matches 2596; Conservative 0; Mismatches 177; Indels 609; Gaps 12;
  QY 1 GTCGAGAGGGCGGCAGTGAGGAGGTTACCTACCTCCGTCCTCAAGGTAAAGGAGCAGTACC 60
    |||
  DB 147635 GTCGAGAGGGGTGGCGGTGAGGAGA----TACCCCTCGTCCAAAGGTAAAGGAGCAGGCAC 147689
    |||
  QY 61 TGCAGCTTTGCTGGAGCAGCCGCTAAAGAGATACCCCGCCCAAGGTAAAGAGAAACCCCAAG 120
    |||

```



```

Db 150982 CATGACTGAGCGACTGAACTGA 151003
|||||
AC147195 223761 bp DNA linear HTG 17-DEC-2003
Bos taurus clone RP42-466M15, WORKING DRAFT SEQUENCE, 4 ordered
pieces.
AC147195
AC147195.2 Gi:39979451
HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 223761)
Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
Idol,J.R., Jones,C., Karlins,E., Kim,H., Kwong,P., Laric,P.,
Larson,S., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Marquies,E.H., Masello,C., Maskeri,B., McDowell,J.,
Mullikin,J.C., Pagnirigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
Reddik-Dugue,N., Schandler,K., Schueler,M.G., Shah,K., Sison,C.,
Stantipop,S., Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L.,
Wetherby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
REFERENCE
2 (bases 1 to 223761)
Green,E.D.
Direct Submission
Submitted (08-NOV-2003) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 223761)
Green,E.D.
Direct Submission
Submitted (17-DEC-2003) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
On Dec 17, 2003 this sequence version replaced gi:38194246.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hghri.nih.gov
----- Project Information
Center project name: ees
Center clone name: 466M15

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 223320 bases at least Q40
Consensus quality: 223438 bases at least Q30
Consensus quality: 223450 bases at least Q20
Insert size: 178000; agarose-fp
Insert size: 223461; sum-of-contigs
Quality coverage: 27.55x in Q20 bases; agarose-fp

```

 Quality coverage: 21.94x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 145383: contig of 145383 bp in length
 * 145384 171206: contig of 25723 bp in length
 * 145484 171306: gap of unknown length
 * 171207 209669: contig of 38363 bp in length
 * 171307 209670: gap of unknown length
 * 209670 223761: contig of 13992 bp in length.
 * 209770
 Location/Qualifiers
 1. 223761
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-466M15"
 /clone_lib="RP42"
 /note="BAC resource: http://bacpac.chori.org/"
 misc_feature
 1. 145383
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left
 145484..171206
 /note="assembly_fragment"
 171307..209669
 /note="assembly_fragment"
 209770..223761
 /note="assembly_fragment"
 clone_end:17
 vector_side:right

Query Match 57.8%; Score 2065; DB 1; Length 223761;
 Best Local Similarity 76.7%; Pred. No. 0;
 Matches 2561; Conservative 0; Mismatches 165; Indels 614; Gaps 9;

QY 81 GTAAGAGATACCCACGCCCAAGGTAGAGAGAAACCCAGTAAGATGTTGTGA 140
 Db 198857 GTGAGGAATATCCCGTCCTCCAAAGTAAGAGAAACCCAGTAGGTGTGCAA 198916
 QY 141 GAGGGCATCAGAGGGCAGACATCTGAAACCATACACGAGAAACCTAGTCAATCTAATC 200
 Db 198917 GAGGGCATCAGAGGGCAGACACACCTGAAACCATCTCAGAGAACTAGTCAATCTAATC 198976
 QY 201 ACATAGGACACAGCCTTGTCTAATCAATGAATCAAGCATCCCGTGGGCAACCC 260
 Db 198977 ACATAGGACACAGCCTTGTCTAATCAATGAATCAAGCATCCCGTGGGCAACCC 199036
 QY 261 AGATGGGCGAGTCATGGTGGAGAGATCTGACAGATGTGTCCACTGGAGAGGAAT- 319
 Db 199037 AGACGGGCGAGTCATGGTGGAGAGAACTGACAGATGTGTGTCTACTGGAGAGGAATG 199096
 QY 320 GCAAACCACTTCAGTATTCTTGCCTTGAGAACCCCATGAACAGTATGAAAGGCGAAATG 379
 Db 199097 GCAAACCACTTCAGTATTCTTGCCTTGAGAAATCC-ATGAAAGGCGAAATG 199147
 QY 380 ATAGATCTGAAGAGGAACTCCCGAGTGTGAGTGGCCCATCTACTCTGAGATC 439
 Db 199148 ATAGATCTGAAGAGGAACTCCCGAGTGTGAGTGGCCCATCTACTCTGAGATC 199207
 QY 440 AGTGGAGAAATTAATCCGAAAGAAATGAAGAGATGAGGCCAAAGCAAAAGAAATCCAG 499
 Db 199208 AGTGGAGAAATTAATCCGAAAGAAATGAAG-GACGAGCCCAAGCAAAACAAATACCCAA 199266
 QY 500 CTGTGATGTGACTGGTGATATAACCAAGTCCGATGCTGTAAAGAGCAATATTCATAG 559

Db	199267	CTGTGGATGTGACTGGTAAATAGAGCAAGGTCTGATGCTCTTAAGAGCAATATATTGCATAG	199326
Qy	560	GAACCTGGAATGTGTCAGGTCCATGAATCAAGGCCAAATTTGGAAGTGGTCAAAACAAGAGATGG	619
Db	199327	GAACCTGGAATGTGTCAGGTCCATGAATCAAGGCCAAATTTGGAAGTGGTCAAAACAAGAGATGG	199386
Qy	620	CNAGAGTGAATGTCAACATTTCTAGGAATCAGCAATCAATTAATGGAATGGTGGTGAAT	679
Db	199387	GAAGAGTGAATGTGCAATTTCTAGGAATCAGCAATCAATTAATGGAATGGTGGTGAAT	199446
Qy	680	TTAACTCAGATGACCAATATATCTACTCTGCGGCGAGGAATCCCTCAGAGAAATGGAG	739
Db	199447	TTAACTCAGATGACCAATATATCTACTCTGCGGCGAGGAATCCCTCAGAGAAATGGAG	199506
Qy	740	TAGCCATCTAGTCCAAACAAGAGTCCGAAATGCAAGTACTTTGATGCAATCTCAAAACG	799
Db	199507	TGGCCATCTAGTCCAAACAAGAGTCCGAAATGCAAGTACTTTGATGCAATCTCAAAACG	199566
Qy	800	ACAGAATGATCTCTGTTTGTTCACAGGCCAAACCATTAATCAATATCAAGTAATCCAGTCT	859
Db	199567	ACAGAATGATCTCTGTTTGTTCACAGGCCAAACCATTAATCAATATCAAGTAATCCAGTCT	199626
Qy	860	ATGCCCAACAGGAAATGCTGGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACTCAAG	919
Db	199627	ATGCCCAACAGGAAATGCTGGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACTCAAG	199686
Qy	920	ACCTTTTAGAAGTAAACACCAAAAGATGCTCTCTCTCATTTAGGGGACTGGAATGCAA	979
Db	199687	ACCTTTTAGAAGTAAACACCAAAAGATGCTCTCTCTCATTTAGGGGACTGGAATGCAA	199746
Qy	980	AAGTAGGAAGCAAAAGAACCTTGGAGTAAACAGGCAAAATTTGGCTCTGAATACGGAATG	1039
Db	199747	AAGTAGGAAGCAAAAGAACCTTGGAGTAAACAGGCAAAATTTGGCTCTGAATACGGAATG	199806
Qy	1040	AAGCAGGCGCAAGACTAATAGAGTTTCCCAAGAAATGACCTGGTCTATAGCAACACCC	1099
Db	199807	AAGCAGGCGCAAGACTAATAGAGTTTCCCAAGAAATGACCTGGTCTATAGCAACACCC	199866
Qy	1100	TCCTCCAAACACCAAGAGAGACTCTACATGGACATCACAGATGGTCAACACCGAA	1159
Db	199867	TCCTCCAAACACCAAGAGAGACTCTACATGGACATCACAGATGGTCAACACCGAA	199926
Qy	1160	ATCAGATGATATATTTCTTGGAGCAAGATGGAGAGCTCTATACGTACAGTCAAGCAAAA	1219
Db	199927	ATCAGATGATATATTTCTTGGAGCAAGATGGAGAGCTCTATACGTACAGTCAAGCAAAA	199986
Qy	1220	CAAGCAGGAGCTTACTGTGGCTCAGATCATGCAATTCCTTATGGCCAAATTCAGACTTA	1279
Db	199987	CAAGCAGGAGCTTACTGTGGCTCAGATCATGCAATTCCTTATGGCCAAATTCAGACTTA	200046
Qy	1280	AATTGAAAGAAAGTAGGGAACCACTAGATCACTCAGSTAAGACCTAAATCCCAATCCCTT	1339
Db	200047	AATTGAAAGAAAGTAGGGAACCACTAGATCACTCAGSTAAGACCTAAATCCCAATCCCTT	200106
Qy	1340	ATGATTTACAGTGGAGTGAATAGATTTAAGGCGCTAGATCTGATA-----GACAG	1394
Db	200107	ATGATTTACAGTGGAGTGAATAGATTTAAGGCGCTAGATCTGATA-----GACAG	200166
Qy	1395	AGTACCTAATGAACTATGGACAGAGGTTCTATGACATTTGTAAGAGACAGGGAATCCAGAC	1454
Db	200167	AGTACCTAATGAACTATGGACAGAGGTTCTATGACATTTGTAAGAGACAGGGAATCCAGAC	200226
Qy	1455	CATCCCATGAAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGGAGGCGCTTACAAAT	1514
Db	200227	CATCCCATGAAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGGAGGCGCTTACAAAT	200286
Qy	1515	AGCTGTGAAGAGAGAGAGTGAAGACCAAGGAAAGAAAGAGAGATATAAGAGATCTGAA	1574
Db	200287	AGCTGTGAAGAGAGAGAGTGAAGACCAAGGAAAGAAAGAGAGATATAAGAGATCTGAA	200346
Qy	1575	TGCAGAGTTCAAA-----	1588
Db	200347	TGCAGAGTTCAAAAGAAATAGCAAGAGAGATAAGAAAGCGCTTCTTCAGCGATCAATGCAA	200406
Qy	1589	-----	1589
Db	200407	ACAAATAGAGGAAAAACAACAGAAATGAGAAAGACTAGAGATCTCTTCAAGAAAAATCAAGAT	200466
Qy	1589	-----	1588
Db	200467	ACCAAAGGAACATTTCTATGCAAAAGATGGCTTGATTAAGGACAGAAATGGTATGACCTA	200526
Qy	1589	-----	1588
Db	200527	ACAGAAGCAGAAGATATTAAAGAAGAGATGGCAATAATACACAGAGAACTGTACAAAAA	200586
Qy	1589	-----	1588
Db	200587	GATCTTCAACCCAGATAATCACGATGGTGTGATCACTGACCTAGAGCCAGACATCTCTG	200646
Qy	1589	-----	1588
Db	200647	GAATGTGAAGTCAAGTGGGCGCTTAGAAAGCATCACTACGAACAAAGCTAGTGGAGGTGAT	200706
Qy	1589	-----	1588
Db	200707	GGATTTCCAGTTGAGCTATTCCAAATCCCTGAAAGATGATGCTGTGAAAAGTCTGCACTCA	200766
Qy	1589	-----	1588
Db	200767	ATATGCCACACATTTTGGNAAACTCAGCAGTGGCCACAGGACTAGAAAAGGTCACTTTT	200826
Qy	1589	-----	1588
Db	200827	ATTCCAATCCCAAGAAAGGCAACGCCAAGAAATGCTCAAACTACTGCACAAATTCGACTC	200886
Qy	1589	-----	1588
Db	200887	ATCTCACACACTAGTAAGTAATGCTCAACATTTCCAAAGCCAGGCTTCAGCAATATCTG	200946
Qy	1589	-----	1642
Db	200947	AACCTGAACTTCTGATGCTTCAAGTGGTTTAGAAAGTCCAGAGAACCCAGAGCCAA	201006
Qy	1643	ATTGCCACATCTCTGTATCATGAAAAGCAAGAGAGTCCAGAAAAACATCTATTTC	1702
Db	201007	ATTGCCACATCTCTGTATCATGAAAAGCAAGAGAGTCCAGAAAAACATCTATTTC	201066
Qy	1703	TGCTTTATTGACTATGCAAAAGCCTTTGACTGTGGGGTCAAAATAAACTGTGGAAAAAT	1762
Db	201067	TGCTTTATTGACTATGCAAAAGCCTTTGACTGTGGGGTCAAAATAAACTGTGGAAAAAT	201126
Qy	1763	CTGAAAGGATGGGAATACACAGACCTGACCTGACTCTTGAAAAATTTGTATGAGGT	1822
Db	201127	CTGAAAGGATGGGAATACACAGACCTGACCTGACTCTTGAAAAATTTGTATGAGGT	201186
Qy	1823	CAGGAAGCAACAGTTAGAACTGACATGGAACCAACAGACTGGTTCCTCAAGTAGGAAAAAG	1882
Db	201187	CAGGAAGCAACAGTTAGAACTGACATGGAACCAACAGACTGGTTCCTCAAGTAGGAAAAAG	201246
Qy	1883	GTATGTCAGGCTGTATATTGTCACCCGCTTTGTTAACTTCTATGCAAG-ACATCATG	1941
Db	201247	GTATGTCAGGCTGTATATTGTCACCCGCTTTGTTAACTTCTATGCAAG-ACATCATG	201306
Qy	1942	AGAAACGCTGGCTGGAGAGAACCAAGCTGGAATCAAGNTTCCGGGAGGAATAGCAAT	2001
Db	201307	AGAAACGCTGGCTGGAGAGAACCAAGCTGGAATCAAGNTTCCGGGAGGAATAGCAAT	201366
Qy	2002	AACCTCAGATATGCAAGTATACCAACCTTATGGCAGAAAGTGAAGAGAACTAAAAAGC	2061
Db	201367	AACCTCAGATATGCAAGTATACCAACCTTATGGCAGAAAGTGAAGAGAACTAAAAAGC	201426
Qy	2062	CTCTTATGAAGGTGAAGAGAGAGTGAAGAGTGGCTTAAAGCTCAACATTCAGAAA	2121
Db	201427	CTCTTATGAAGGTGAAGAGAGAGTGAAGAGTGGCTTAAAGCTCAACATTCAGAAA	201486

Db	44301	GGACCTAA	CAGAAGCAGAAGATATTAAAGAGAGGTGGCAAGAAATACACAGAACTGTACAA	44360
Qy	1589	-----	-----	1588
Db	44361	AAGAATCTT	CACGACCCAGATAATCAGATGGTGTGATCACTGACCTAGACCCAGACAT	44420
Qy	1589	-----	-----	1588
Db	44421	CCTGGAATGTGAAGTCAAGTGGGCGCTTAGAAAGCATCTAGGAACAAGCTAGTGGAGG	44480	
Qy	1589	-----	-----	1588
Db	44481	TGATGGAATTC	CAGTTGAACTTTTCCAAATCCTGAAAGATGATGCTGTGAAAGTGTGCA	44540
Qy	1589	-----	-----	1588
Db	44541	CTCAATATGCCAGCAAA	TTTGGAAACTCAGCAGTGGTACAGGACTGGAAAAGTCACT	44600
Qy	1589	-----	-----	1588
Db	44601	TTTTATTCCAAATCCCAAGAAAGCAATGCCAAAGATGCTCAAACTACCAACAATCAC	44660	
Qy	1589	-----	-----	1588
Db	44661	ACTCATCTCGCAGCTAGTCAAGTAATGCTCAAAATTTCCCAAGCCAGGCTTCAGCAATA	44720	
Qy	1589	-----	-----	1638
Db	44721	TGTGAAC	TGTGAACTTCTCTGATGTTCAAGCTGGTTTAGAAAAGCAAGAAACCAAGACA	44780
Qy	1639	CCAAATTC	CAACATCCTCTGTATCATGGAAGAAACCAAGAGAGTTCACAGAAAACATCTA	1698
Db	44781	TCAAATTC	CAACATCCGCTGGATCATCGAAAAGCAAGAGAGTTCACAGAAAACATCTA	44840
Qy	1699	TTTTCTGCTTTAT	TGACTATGCAAAAGCCTTTGACTGTGGGGTCAACAATAAATCTGTGAA	1758
Db	44841	TTTTCTC	TTTATTGACTATGCCAAAGCCTTTGACTGTGTGGATCAAAATAAATCTGTGAA	44900
Qy	1759	AATCTG	AAAGGATGGGAATACCAAGACCACTGACCTGACTCTTTGAAAATTTGTATGC	1818
Db	44901	AATCTG	AAAGAGATGAGAATACCAAGACCACTGATCTGCTCTTTGAGAAATTTGTATGC	44960
Qy	1819	AGGTCAGGAAGCA	ACAGTTAGAACTGGACATGGAAACAACAGACTGGTTCCAAAGTAGGAAA	1878
Db	44961	AGGTCAGGAAGCA	ACAGTTAGAACTGGACATGGAAACAACAGACTGGTTCCAAATAGAAAA	45020
Qy	1879	AGAGTATG	CAAGGCTGTATATGTCAACCGCTTGTTTAACTTCTATGACAG-ACAT	1937
Db	45021	AGGAGTACGT	CAAGGCTGTATATGTCAACCTTGTATGTCAACCTTGTATATGACAGAGTACAT	45080
Qy	1938	CATGAGAAACGT	GGCTGGAGAAGCAACAGCTGGAATCAAGATTGCGGGAGAAAATAG	1997
Db	45081	CATGAGAAACGT	GGCTGGAGAAGCAACAGCTGGAATCAAGATTGCGGGAGAAAATAT	45140
Qy	1998	CAATAACCT	CAGATATGACAGATGATACCAACCTTATGGCAGAAAAGTGAAGAGGAACCTAAA	2057
Db	45141	CAGTAAACCT	CAGATATGATGATGACACCAACCTTATGGCAGAAAAGTGAAGAGGAACCTAAA	45200
Qy	2058	AAGCCTCTT	GATGAAGGTGAAGAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCA	2117
Db	45201	AAGCCTCTT	GATGAAGGTGAAGTGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCA	45260
Qy	2118	GAATAAC	GAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAATAGATGGGGAACAGT	2177
Db	45261	GAATAAC	GAAGATCATGGCGCCGGTCCCAACCTTCATGGGAAATAGATGGGGAACAGT	45320
Qy	2178	GGAAACAGT	GTACAGCTTTATTTTGGGGGCTCCAAATCACTGCAGATGGTGACTGC	2237
Db	45321	GGAAACAGT	GTACAGCTTTATTTTTTCTGGGCTCCAAATCACTGCAGATGGTGACTGC	45378
Qy	2238	AGCCATG	GAATTAATAAGACACTTACTCTTGTGAAGAAAAGTTTA--ACCAACTAGATAGC	2295

[illegible]

JOURNAL Submitted (24-SEP-2003) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
COMMENT On Sep 24, 2003 this sequence version replaced gi:34328552.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoonhgri.nih.gov
----- Project Information
Center project name: eed
Center clone name: 234122

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 184412 bases at least Q40
Consensus quality: 184761 bases at least Q30
Consensus quality: 184992 bases at least Q20
Insert size: 186000; agarose-fp
Insert size: 185030; sum-of-contigs
Quality coverage: 17.20x in Q20 bases; agarose-fp
Quality coverage: 17.29x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.
* 1 53583: contig of 53583 bp in length
* 53584 53683: gap of unknown length
* 53684 54553: contig of 3770 bp in length
* 54554 57553: gap of unknown length
* 57554 83676: contig of 26123 bp in length
* 83677 83776: gap of unknown length
* 83777 88070: contig of 4294 bp in length
* 88071 88170: gap of unknown length
* 88171 175702: contig of 87532 bp in length
* 175703 175802: gap of unknown length
* 175803 185590: contig of 9788 bp in length.
Location/Qualifiers
1..185590
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-234122"
/clone_lib="RP42"

FEATURES
source
1..185590
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-234122"
/clone_lib="RP42"

misc_feature
1..53583
/note="assembly_fragment"
vector side:left

misc_feature
1..26399
/note="clone overlaps with GenBank Accession Number. AC144997 clone CH240-434022 (center project name eeh)"

misc_feature
53684..57453
/note="assembly_fragment"

misc_feature
57554..83676

/note="assembly_fragment"
83777..88070
/note="assembly_fragment"
88171..175702
/note="assembly_fragment"
136271..185590
/note="clone overlaps with GenBank Accession Number. AC145843 clone RP42-230A24 (center project name eec)"
175803..185590
/note="assembly_fragment"
clone_end:SP6
vector_side:right

Query Match 57.5%; Score 2052.6; DB 1; Length 185590;
Best Local Similarity 75.0%; Pred. No. 0;
Matches 2624; Conservative 0; Mismatches 239; Indels 636; Gaps 11;

QY 1 GTCAGGAAGGGCGGAGTGAGGAGGAGTACCTCCCTGTCACAGGTAAAGGAGGAGTACG 60
Db 36034 GTCAGGAAGGGCGGCGGTGAGGAGA-----TACCCTCTGTCACAGGTAAAGGAGGAGTACG 36088
QY 61 TCGGCTTTCTCGAGCAGCCGTAAGAGATATCCACGCCCAAGGTAAAGAGAAACCCCAAG 120
Db 36089 TGTGCTTTCTCGAGCAGCCGTAAGAGATATCCACGCCCAAGGTAAAGAGAAACCCCAAG 36148
QY 121 TAAGATGTTAGTGTGTTGAGAGGGCATCAGAGGGCAGACATATCTGAAACCATACACGCA 180
Db 36149 TAAGACAGTAGTGTGTTGCAAGAGGGCATCAGA-AGCAAAACACACTTAAACCATACTCACA 36207
QY 181 GAAACTAGTCAATCTAATCAGTACACTAGGACACAGCCTGTCTAATCAATCAAACTAAG 240
Db 36208 GAAACTAGTCAATCTAATCAGTACACTAGGACACAGCCTGTCTAATCAATCAAACTAAG 36267
QY 241 CCATGCCCTGGGGCAACCCCAAGATGGGAGGTCTATGTTGGAGAGATCTGCAGAGATGTG 300
Db 36268 CCATGCCCTGGGGCAACCCCAAGATGGGAGGTCTATGTTGGAGAGATTTGACAGATGTG 36327
QY 301 GTCCACTGAGAGAGGAAT-GCAACCACTCAGTATTTCTGCTGTGAGAAACCCCATGAA 359
Db 36328 GTCCACTGAGAGAGGAATGGCAACCACTCAGTATTTCTGCTGTGAGAAACCCCATGAA 36387
QY 360 CAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACTCCCGAGTCACTAGGTGC 419
Db 36389 CAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAACTCCCGAGTCACTAGGTGC 36447
QY 420 CCATATGCTTCTGAGATCAGTGGAGAAATAACTCCAGAAAGATGAAGAGATGAGGCC 479
Db 36448 CCAATATGCTTCTAGAGATCAGTGGAGAAATAACTCCAGAAAGATGAAGGATGAAGCC 36507
QY 480 AAAGCAAAAGATACCCAGCTGTGATGCTGCTGTGATATAAGCAAGSTCCGATGCTG 539
Db 36508 AAAGCAAAAGATACCCAGCTGTGATGCTGCTGTGATATAAGCAAGSTCCGATGCTG 36567
QY 540 TAAAGAGCAATATTGCATAGGAACCTGGAAATGTCCAGTCCATGATCAAGGCAAAATGGA 599
Db 36568 TAAAGAGCAATATTGCATAGGAACCTGGAAATGTCCAGTCCATGATCAAGGCAAAATGGA 36627
QY 600 AGTGGTCAAAAGAGATGCAAGAGATGAATGTCAACATTTAGGAAATCAGGAACTTAA 659
Db 36628 AGTGGTCAAAAGAGATGCAAGAGATGAATGTCAACATTTAGGAAATCAGGAACTTAA 36687
QY 660 ATGGACTGGAAATGGTGAATTTAACTCAGATGACCATTTATCTACTTCTGCGGCGAGGA 719
Db 36688 ATGGACTGGAAATGGTGAATTTAACTCAGATGACCATTTATCTACTTCTGCGGCGAGGA 36747
QY 720 ATCCCTCAGAAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCGAAATCGACTACT 779
Db 36748 ATCCCTCAGAAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCGAAATCGACTACT 36807
QY 780 TGAATGCACTCTCAAAAGAGAGATGATCTCTGTTGTTTCCAGGCAAAACCATTCAA 839
Db 36808 TGAATGCACTCTCAAAAGAGAGATGATCTCTGTTGTTTCCAGGCAAAACCATTCAA 36867

Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoonhgrl.nih.gov
----- Project Information
Center project name: crk
Center clone name: 056C13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 123475 bases at least Q40
Consensus quality: 123571 bases at least Q30
Consensus quality: 123601 bases at least Q20
Insert size: 111000; agarose-fp
Quality coverage: 10.67X in Q20 bases; agarose-fp
Quality coverage: 9.58X in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 35692: contig of 35692 bp in length
* 35693 123710: gap of unknown length
* 35793 123710: contig of 87918 bp in length.

Location/Qualifiers
1. 123710
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-56C13"
/clone_lib="RP42"
/notes="assembly_fragment
clone end:17
vector side:left"
35793_123710
/notes="assembly_fragment
clone end:SP6
vector side:right"
111382_123710
/notes="clone overlaps with GenBank Accession Number
AC116433 clone RP42-556E7 (center project name crk)"

Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoonhgrl.nih.gov
----- Project Information
Center project name: crk
Center clone name: 056C13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 123475 bases at least Q40
Consensus quality: 123571 bases at least Q30
Consensus quality: 123601 bases at least Q20
Insert size: 111000; agarose-fp
Quality coverage: 10.67X in Q20 bases; agarose-fp
Quality coverage: 9.58X in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 35692: contig of 35692 bp in length
* 35693 123710: gap of unknown length
* 35793 123710: contig of 87918 bp in length.

Location/Qualifiers
1. 123710
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-56C13"
/clone_lib="RP42"
/notes="assembly_fragment
clone end:17
vector side:left"
35793_123710
/notes="assembly_fragment
clone end:SP6
vector side:right"
111382_123710
/notes="clone overlaps with GenBank Accession Number
AC116433 clone RP42-556E7 (center project name crk)"

Query Match 57.4%; Score 2048.6; DB 1; Length 123710;
Best Local Similarity 76.2%; Pred. No. 0;
Matches 2592; Conservative 0; Mismatches 194; Indels 614; Gaps 12;
QY 1 GTCAGGAAGGGCGGCGAGTGGAGAGGTACCTCGTCCAAAGTAAGGACGAGTACG 60
DB 109448 GTGAGGAGGGCGGTCTGTGAGGAGA-----TACCCCTCTGCCAAGGAGGACGAGTGGC 109502
QY 61 TGCCTTTTGTGAGGAGCGCGTTAAAGATACCCACGCGCAAGGTAAAGAAACCCAG 120
----- Genome Center

Qy	1687	GAAAAACATCTATTTCCTGCTTTTATAGCATATGCAAAAGCCTTTGACTGTGGGGTCAAA	1746
Db	111719	GAAAAACATCTATTTCCTGCTTTTATAGCATATGCAAAAGCCTTTGACTGTGGGTACAA	111778
Qy	1747	TAAACTGTGGAANAATTCGMAAAGGATGGGAATACCAGACCACCTTGAC--CTGACTCTT	1803
Db	111779	TAAACTGTGGAANAATTCGMAAGAGATGGGAATACCAGACCACCTTGACTTCTTCCCTCTT	111838
Qy	1804	GAAAAATTTGTATCGAGGTGAGGAAGCAACAGTTAGAACTGGACATGGAACAACAGACTG	1863
Db	111839	GAGAAACCTTATATCGAGGTGAGGAAGCAACAGTTAGAACTGGACATGGAACAACAGACTG	111898
Qy	1864	GTTCCTCAAGTAGAAAAAGGAGTATGTCAAGGCTGTATATTGTCAACCGGCTTGTTTAACCT	1923
Db	111899	GTTCCTCAATAGAAAAAGGAGTACATCAAGGCTGTGTATTGTCAACCTTGTTTAACTT	111958
Qy	1924	CTATGCAAG--ACATCATGAAAAACGCTGGGCTGGAAGAAGCACAAAGCTGGAATCAAGAT	1982
Db	111959	CTATGCAAGATPACATCATGAAAAACCTGGGCTGGAAGAAGCACAAAGCTGGAATCAAGAT	112018
Qy	1983	TGCCGGAGAAATGCAATAACCTCAGATATGCAAGATGATACCAACCTTATGGCAGAAG	2042
Db	112019	TGCCGGAGAAATTCATAAACCTCAGATATACAGATGACACCAACCTTATGGCAGAAG	112078
Qy	2043	TGAAGAGGAACATAAAAAACGCTCTTTGATGAAGGTGAAAGAGGAGTGAANAAGTTGGCTT	2102
Db	112079	TGAAGAGGAACATAAAAAACGCTCTTTGATGAAGGTGAAAGAGGAGTGAANAAGTTGGCTT	112138
Qy	2103	AAAGCTCAAACATTCAGAAAAAGAGATCATGGCATCTGGTCCCACACTTCATGGAAT	2162
Db	112139	AAAGCTCAAACATTCAGAAAAATTAAGATCATGGCATCTGGTCCCACACTTCATGGAAT	112198
Qy	2163	AGATGGGGAACAGTGGAAAAACAGTGTGAGCTTTATTTTTGGGGGGCTCCAAATCACT	2222
Db	112199	AGATGGGGAACAGTGCNAAACAGTGTCCGACTTTA--TTTTGGGGGGCTACAAATCACT	112256
Qy	2223	GCA--GATGTGATGTCAGCCATGAAATATAAAGACACTTACTCTCTTGGAGAAAGTTA	2280
Db	112257	GCAATCACTGTGTATGTCAGCCATGAAATATAAAGACACTTACTCTCTTGGAGAAAGTTA	112316
Qy	2281	--ACCAACCTAGATAGCATATTTGAAAGCAGAGACATACCTTGGCCAACAAGCCCCATC	2338
Db	112317	TGACCAACCTAGATAGCATATTTAAGACAGAGACACTATTTTGGCCAACAAGTCTGTC	112376
Qy	2339	TAGTC--AAGGCTATGGTTTTTCCAGTGTGATGTATGGATGTGAGTTGGACTGTGAAG	2397
Db	112377	TAGTCAAAAGGCTATGGTTTTTCCGTGTGATGTATGGATGTGAGTTGGACTGTGAAG	112436
Qy	2398	AAAGCTGAGCACTGNAGAAATGATGCTTTTGNACTGTGGTGTGAGAGAGACTCTTGAGA	2457
Db	112437	AAAGCTGAGTGTCTGAAGAAATGATGCTTTTGAACGTGTGGTGTGAGAGAGACTCTTGAGA	112496
Qy	2458	GTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATCTCTGAAGGAGATCAGCCCTGGGATT	2517
Db	112497	GTCCCTTGAACTGCAAGGAGATCCAAACCGTCCATCTCTGAAGGAGATCAGTCTGGTGT	112556
Qy	2518	TCTTTGGAAGGAATGATGCTAAAGCTGAAAATCTCCAGTACTTTTGGCCACTGTATCAGAGA	2577
Db	112557	TCAITGGAAGGACTGATGCTAAAGCTGAAAATCTCCAGTACTTTTGGCCACTCTCAGAGA	112616
Qy	2578	GCTGACTCACTGGAAGAACCTGTATGCTGGAGGGATTTGGGGCAGGAGGAGAGAGGGA	2637
Db	112617	GTGACTCACTTGGAAAAAGACTCTGTATGCTGGAGGGATTTGGGGCAGGAGGAGAGGGA	112676
Qy	2638	CGACAGAGGATGAGATGGTGGATGGCATCACTGATCGATGGAAGTGTGGGTGAA	2697
Db	112677	TGACAGAGGATGAGATGGTGGATGGCATCACTGATGGAAGTGTGGGTGAA	112736
Qy	2698	CTCCTGGAGTTGGTGAATGGAAGGAGGCTCTCTGGGCGATTCATGGGCTCAAGAG	2757
Db	112737	CTCTGGGAGTTGGTGAATGGAAGGAGGCTCTGGGCTGCTGCCGATTCATGGGCTGCCAAG	112796
Qy	2758	AGTTGGACACGACTGAGCAACTGAACTGAACTGAACTGTA	2797

Db 112797 AGTCGACATGACTAGTACTGACTGACTGACTGAA 112836

RESULT 26
AC091399 linear HTG 01-JUN-2002
LOCUS Bos taurus clone RP42-65K24, WORKING DRAFT SEQUENCE, 5 ordered
DEFINITION pieces.
AC091399
VERSION AC091399.2 GI:21306657
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 181921)
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
Lee-lin,S.-O., Legaspi,R., Maduro,O.L., Maduro,V.B.,
Margulis,E.H., Masiello,C., Maskeri,B., Masrrian,S.D.,
McCloskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Stancirip,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurgon,C., Vogt,J.L., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 181921)
Green,E.D.
Direct Submission
Submitted (19-APR-2001) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 181921)
Green,E.D.
Direct Submission
Submitted (01-JUN-2002) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
On Jun 1, 2002 this sequence version replaced gi:13677070.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: axb
Center clone name: 065K24

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid, n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 180679 bases at least Q40
Consensus quality: 181147 bases at least Q30
Consensus quality: 181409 bases at least Q20
Insert size: 178000; agarose-fp
Insert size: 181821; sum-of-contigs
Quality coverage: 11.36x in Q20 bases; agarose-fp
Quality coverage: 11.14x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 27632: contig of 27632 bp in length
 * 27633 27732: gap of unknown length
 * 27733 69608: contig of 41876 bp in length
 * 69609 69708: gap of unknown length
 * 69709 109185: contig of 39477 bp in length
 * 109186 109285: gap of unknown length
 * 109286 181783: contig of 72498 bp in length
 * 181784 181883: gap of unknown length
 * 181884 181921: contig of 38 bp in length.

FEATURES

source

1. 181921
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-65K24"
 /clone_lib="RP42"

misc_feature

1. 75601
 /note="clone overlaps with GenBank Accession Number
 AC091660 clone RP42-354B6 (center project name axa)"

misc_feature

1. 27632
 /note="assembly_fragment
 clone end: SP6
 vector_side:left"

misc_feature

27733..69608
 /note="assembly_fragment"
 69709..109185
 /note="assembly_fragment"

misc_feature

109286..181783
 /note="assembly_fragment"
 181884..181921
 /note="assembly_fragment
 clone end: T7
 vector_side:right"

misc_feature

Query Match 57.3%; Score 2046.4; DB 1; Length 181921;
 Best Local Similarity 74.1%; Pred No. 0;
 Matches 2592; Conservative 0; Mismatches 201; Indels 706; Gaps 9;

QY	1	GTCCAGGAGGGCGGAGTGAGAGAGGTACTACCTCGTCCAAAGGTAAAGAGCAGTACG	60
DB	66540	GTCCAGGAGGGCGGCGGTGAGGAGA-----CACCCCTCGTCCAAAGGTAAAGAGCAGTACG	66594
QY	61	TGCGCTTGCTGGAGCAGCCCTAAGAGATACCCAGCCCGCCAGGTAAAGAGAACCCAG	120
DB	66595	TGCGCTTGCTGGAGCAGCCCTAAGAGATACCCAGCCCGCCAGGTAAAGAGGTACCGAAG	66654
QY	121	TAAGATGTTAGTGTGTGTGAGAGGGCATCAGAGGGCGAGACATCTGAAACCATACAGCA	180
DB	66655	TAAGACGGTAGTGTGTGCAAGAGGGCATCAGAGGGCGAGACATCTGAAACCATACAT	66714
QY	181	GAAACTAGTCAATCTAATCACTAGACACAGCCCTGTCTTAACCTCAATGAACTAAG	240
DB	66715	GAAACTAGTCAATCTAATCACTAGAAACCCAGCCCTGTCTTAACCTCAATGAACTAAG	66774
QY	241	CCATGCCCGTGGGCAACCCCAAGATGGCAGGTCAATGTTGAGAGATCTGCACAGAATGTG	300
DB	66775	CCATGCCCGTGGGCGCCACCAAGACGGGTGGGTCATGTTGGAGAGGTCTGCACAGAATGTG	66834
QY	301	GTCCACTGGAGAGGGGAAT-GCAAAACCACTTCAGTATTCTTGCCTTGAGAACCCCATGAA	359
DB	66835	GTCCACTGGAGAGGGGAATGCAAAACCACTTCAGTATTCTTGCCTTGAGAACCCCATGAA	66894
QY	360	CAGTATGAAAAGGCAAAATGATAGGATCTGAAAGAGGAACCTCCCGGTCAGTAGTGTC	419

DB	66895	CAGTATGAAAAGGCAAAATGATAGGATACAGAAAGAGGAATCCCGAGGTCAAGTAGTGTC	66954
QY	420	CCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGAGCC	479
DB	66955	CCAGTATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGAGCC	67014
QY	480	AAAGCAAAAAGAAATACCCAGCTGTGGATGTGATGTGATATATAAGCAAGGTCCGATGCTG	539
DB	67015	AAAGCAAAAAGAAATACCCAGCTGTGGATGTGATGTGATATATAAGCAAGGTCCGATGCTG	67074
QY	540	TAAAGAGCAATATTCGATAGGAACCTGGATGTCAGGTCCATGATCAAGGCAAAATGGA	599
DB	67075	TAAAGAGCAATATTCGATAGGAACCTGGATGTCAGGTCCATGATCAAGGCAAAATGGA	67134
QY	600	AGTGGTCAAAACAGAGATGCGCAAGAGTGAATCTCAACATCTTAGGAATCAGCAAACTAAA	659
DB	67135	AGTGGTCAAAACAGAGATGCGCAAGAGTGAATCTCGACATCTTAGGAATCAGCAAACTAAA	67194
QY	660	ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTCTGCGGCAGGA	719
DB	67195	ATGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTCTGCGGCAGGA	67254
QY	720	ATCCCTCAGAGAAATGGAGTAGCCATCATGTCTCAACAAAGAGTCCGAAATGCAAGTACT	779
DB	67255	ATCCCTTAGAGAAACGGAGTAGCCCTTGTACAGAAACAAAGAGTCCAAATGCAAGTACT	67314
QY	780	TGATGCACTCTCAAAAACGACAGAAATGATCTCTGTTTCCAAAGGCAAAACCAATCAA	839
DB	67315	TGATGCACTCTCAAAAACGACAGAAATGATCTCTGTTTCCAAAGGCAAAACCAATCAA	67374
QY	840	TATCAGCAATATCCAGTCTATGCCCCACAGTAACTGCTGAAGAGCTGAAGTTGAACG	899
DB	67375	TATCAGCAATATCCAGTCTATGCCCCACAGTAACTGCTGAAGAGCTGAAGTTGAACG	67434
QY	900	GTCTATGAAGACCTACAGACCTTTTAGAAGCTAAACCCCAAAAGAGATGCTCTTCTCAT	959
DB	67435	GTCTATGAAGACCTACAGACCTTTTAGAAGCTAAACCCCAAAAGAGATGCTCTTCTCAT	67494
QY	960	TATAGGGAGTGAATGCAAAAGTAGGAAGCAAGAACACCTGAGTAAACAGGCAATTT	1019
DB	67495	TATAGGGAGTGAATGCAAAAGTAGGAAGTCAAGAAACCTGAGTAAACAGGCAATTT	67554
QY	1020	TGCGCTTGGAATACGGAATGAAGAGGCAAGAGCTAATAGAGTTTTCGCAAGAAATGTC	1079
DB	67555	TGCGCTTGGAATACGGAATGAAGAGGCAAGAGCTAATAGAGTTTTCGCAAGAAATGTC	67614
QY	1080	ACTGTCATAGCAAAACACCTCTTCCAAACACAGAGAGAGACTCTACATGAGACATC	1139
DB	67615	ACTGTCATAGCAAAACACCTCTTCCAAACACAGAGAGAGACTCTACATGAGACATC	67674
QY	1140	ACCATGTCATCAACCCGAAATCAGATTGATTATATCTTTTCGACCCAAAGATGAGAGAG	1199
DB	67675	ACCATGTCATCAACCCGAAATCAGATTGATTATATCTTTTCGACCCAAAGATGAGAGAG	67734
QY	1200	CTCTATACAGTCAGCAAAAACCAAGACCCAGGAGCTTACTGTGCTCAGATCATGAACTCCT	1259
DB	67735	CTCTATACAGTCAGCAAAAACCAAGACCCAGGAGCTTACTGTGCTCAGATCATGAACTCCT	67794
QY	1260	TATTCGCAATTCAGCAATTAATTCGAAGAGTAGGGAACCACTAGATCACTCAGGTA	1319
DB	67795	TATTCGCAATTCAGCAATTAATTCGAAGAGTAGGGAACCACTAGATCACTCAGGTA	67854
QY	1320	AGACCTTAATCCAAATCCCTTATGATTATACAGTGAAGTGAAGAAATAGATTAAAGGGCT	1379
DB	67855	AGACCTTAATCCAAATCCCTTATGATTATACAGTGAAGTGAAGAAATAGATTAAAGGGCT	67914
QY	1380	AGATCTGATAGACAGTACCTTAATGAACATATGACAGAGTTTCAATGACATGTACAGGA	1439
DB	67915	AGATCTGATAGACAGTACCTTAATGAACATATGACAGAGTTTCAATGACATGTACAGGA	67974
QY	1440	GACAGGATCGAGACACATCCCATCGAAAAAGAAATGCAAAAAAGCAAAATGGCTGTCTGG	1499

Db 67975 GACAGGATCAAGACCATCTCTTTGGAAAAAGAAATGCAAAAAAGCAAAATGGCTGTCTGG 68034
Qy 1500 GGAGCCCTTACAAATAGCTGTGAAAGAGAGAGAGTGAAGAGCAAGGAAAAAGAGAAAG 1559
Db 68035 GGAGCCCTTACAAATAGCTGTGAAAGAGAGAGAGTGAAGAGCAAGGAAAAAGAGAAAG 68094
Qy 1560 ATAAAGCATCTGAATGCAGAGTTCCAAA 1588
Db 68095 ATATAAGCATCTGAATGCAGAGTTCCAAAAGATAGCAAGAGAGATAAGAAAGCCTTCCT 68154
Qy 1589 1588
Db 68155 CAGTGATCAATGCAAGAAATAGAGGAAATTAACAGAAATGGGAAAGACTAGAGATCTCTT 68214
Qy 1589 1588
Db 68215 CAAGAAATTAGAGATACCAAGGGAACATTTTCATGCAAGAGTGGCTCGATAAAGGACAG 68274
Qy 1589 1588
Db 68275 AAATGGTATGACCTAACAGGAGCAGAGATATTTAAAAAGAGTGGCAGGAATACAGAG 68334
Qy 1589 1588
Db 68335 AGAAGTACAAAAAGAGCTTCAGACCCAGATAATCAGATGGTGTGATCACTGACCT 68394
Qy 1589 1588
Db 68395 AGAGCCAGACATCTGGAATGTGAAGTCAAGTGGGCTTAGAAAGCATCTCTACAAAAA 68454
Qy 1589 1588
Db 68455 AAAAAAAGAAAGCATCTTACGAAACAAAGTAGTGGAGTGATGAATTCAGTGGAG 68514
Qy 1589 1588
Db 68515 CTATTTCAAATCCTGAAAGATGATGCTGTGAAAGTGTGCATCAATATGCCAGAAAT 68574
Qy 1589 1588
Db 68575 TGGAAACTCAGCAGTGGCCACAGGACTGGGAAAGGTGAGTTTCATTCCAATCCCAAAA 68634
Qy 1589 1588
Db 68635 AAAGCAATGCCAAGATGTCTCAACTACTGCACAAATTTGACTCATCTCACATGCTAGT 68694
Qy 1589 1597
Db 68695 AAAGTAATGCTCAAAATTTCTCAAGCTAGGCTTCAGCAATATGTCAACCGTGAACCTTCT 68754
Qy 1598 GTTGTTCAGCTGGTTTGAAGAACTCAGAGGAAACAGAGACCAAAATTTGCCAACATCTC 1657
Db 68755 GATGTTCAAGCTGGTTTGAAGAGGCAAGGAAACAGAGATCAAAATTTGCCAATATCCAC 68814
Qy 1658 TGATCATGAAAAAGCAGAGAGTTCCAGAAAAACATCTATTTCTGCTTTATTGACTAT 1717
Db 68815 TGGATCATCAAAAAAGCAGAGAGTTCCAGAAAAACATCTATTTCTGCTTTATTGACTAT 68874
Qy 1718 GCAAAAGCTTTGACTGTGGGGTCAAAATAAAGTGTGGAATTTCTGAAAGGAGTGGGA 1777
Db 68875 GCCAAAGCCTTTGACTGTGTGATCAAAATAAGTGTGGAATTTCTGAAAGAGATGGGA 68934
Qy 1778 ATACAGACCACTGACCTGACTCTTTGAAAAATTTGTATGCAAGTTCAGGAGTCAAGCAACAGTT 1837
Db 68935 ATACAGACCACTGACCTGCTCTTTGAAACCTATATACAGTTCAGGAGCAACAGTT 68994
Qy 1838 AGAAGTGAATGGAACACAGACTGGTTCCAGTAGGAAAAAGGAGTATGTCAAGGCTGT 1897
Db 68995 AGAAGTGGCCATGGAACACAGACTGGTTCCAAATAGGAAAAAGGAGTACGTCAAGGCTGT 69054
Qy 1898 ATATGTCAACCCGCTGTGTTAACTTCTATGAGAG-ACATCATGAGAAACGCTGGGCTG 1956
Db 69055 ATATGTCAACCTGCTTATTTAACTTATGAGAGTATCATGAGAAACACTGGGCTG 69114

Qy 1957 GAAGAGCAACAGCTGGAATCAAGATTGCCGGAGAAATAGCAATAACCTCAG-ATATGC 2015
Db 69115 GAAGAGCAACAGCTGGAATCAAGATTGCCGGATAAATATCAATAACCTCAGAATATGC 69174
Qy 2016 AGATGATACCACTTATGGCAGAAAGTGAAGAGGAACTTAAAGAGCTCTTCTGATGAAGT 2075
Db 69175 AGATGACACACCTTATGGCAGAAAGTGAAGAGGAACTTAAAGAGCTCTTCTGATGAAGT 69234
Qy 2076 GAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTTCAAGAAACGAAGATCATGCC 2135
Db 69235 GAAGAGGAGAGTGAAGAAATTTGGCTTAAAGCTCAACATTTCAAGAAACGAAGATCATGCC 69294
Qy 2136 ATCTGTCTCCATCACTTATGGGAAATAGATGGGAAACAGTGGAAACAGTGTCAAGCTT 2195
Db 69295 ATCTGTCTCCATCACTTATGGGAAATAGATGGGAAACAGTGGAAACAGTGTCAAGCTT 69354
Qy 2196 TATTTTGGGGGCTCCAAAATCACTGCAGATGGTCACTGCAGCATGAAATTTAAAAAG 2255
Db 69355 TATTTTGGGGCTCCAAAATCACTGCAGATGGTCACTGCAGCATGAAATTTAAAAAG 69412
Qy 2256 CACTTACTCTTGGAGAAAAAGTTA-ACCAACCTAGATAGCATATTTGAAAGCAGAGAC 2313
Db 69413 CGCTTCTCTTGGAGGAAAGTTATGACCAACCTAGATAGCATATTTCAAAAGCAGAGAC 69472
Qy 2314 ATTACCTTCCACAAAGCCCATCTAGTCAAGGCTATGTTTCCAGTGTCTATGAT 2373
Db 69473 ATTACTTTCGCAACAAAGTCCATCTAGTCAAGGCTATGTTTCCAGTGTCTATGAT 69532
Qy 2374 GGATGTGAGAGTTGGACTCTGAAGAAAGCTGAGCACTGAAGAAATTTGATCTTTTGAAGT 2433
Db 69533 GGATGTGAGAGTTGGACTCTGAAGAAAGCTGAGCACTGAAGAAATTTGATCTTTTGAAGT 69592
Qy 2434 TGGTGTGGAAGAGAC----- 2449
Db 69593 TGGTGTGGAAGAGACNN 69652
Qy 2450 -----TCTTGAGAGTCCCTTGGACTGCAAGGAGATCCAA 2483
Db 69653 NNN 69712
Qy 2484 CCAAGTCCATTTGAAAGAGATCAGCCCTGGATTTCTTTGGAAGGAATGATGCTAAAGCT 2543
Db 69713 CCAAGTCCATTTGAAAGAGATCAGCCCTGGATTTCTTTGGAAGGAATGATGCTAAAGCT 69772
Qy 2544 GAAACTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCACTGGAAGAGACCTGAT 2603
Db 69773 GAAACTCCAGTACTTTGGCCACCTGATGAGAGAGTGTGACTCTTGGAAAGAG-CTGAT 69830
Qy 2604 GCTGGAGGAGATTTGGGGCAGGAGAGAGGGGACGACAGAGATGAGATGGCTGATGG 2663
Db 69831 GCTGGAGGAGATTTGGGGCAGGAGAGAGGGGACGACAGAGATGAGATGAATGGATGG 69890
Qy 2664 CATCACTGACTCGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2723
Db 69891 CATCACTGACTCGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 69950
Qy 2724 GGCCTGTCTGCGCGGATTCATGGGCTCAAGAGAGTGGGACACGACTGAGCAACTGAAC 2783
Db 69951 AGCTGGGCTGCTGCGATTCATGGGCTCGAAGAGTGGGACACGACTGAGCGACTGAAC 70010
Qy 2784 TGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2802
Db 70011 TGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 70029

RESULT 27

AC107065
LOCUS AC107065 175949 bp DNA linear NAM 18-SEP-2003
DEFINITION Bos taurus clone rp42-513g13, complete sequence.
ACCESSION AC107065
VERSION AC107065.5 GI:34849950
KEYWORDS HTG.

Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Lalic, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, O.L., Maduro, V.B., Margulies, E.H., Mastello, C., Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schander, K., Schueler, M.G., Sison, C., Stantryp, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 177535)
 Green, E.D.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
 Submitted (08-FEB-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 177535)
 Green, E.D.

Direct Submission
 Submitted (04-DEC-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Dec 4, 2002 this sequence version replaced gi:18640643.
 ----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc_zooenhgri.nih.gov

----- Project Information

Center project name: cbc

Center clone name: 262G19

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 176756 bases at least Q40
 Consensus quality: 177014 bases at least Q30
 Consensus quality: 177134 bases at least Q20
 Insert size: 146000; agarose-fp
 Insert size: 177235; sum-of-contigs
 Quality coverage: 12.93x in Q20 bases; agarose-fp
 Quality coverage: 10.65x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1 25820: contig of 25820 bp in length
 * 25821 25920: gap of unknown length
 * 25921 69500: contig of 44030 bp in length
 * 69511 70050: gap of unknown length
 * 70051 170904: contig of 100854 bp in length
 * 170905 171004: gap of unknown length
 * 171005 177535: contig of 6531 bp in length.

Location/Qualifiers

1. 177535

/organism="Bos taurus"

FEATURES

source

/mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-262G19"
 /clone_lib="RP42"
 1. 25820
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 1. 18394
 /note="clone overlaps with GenBank Accession Number
 AC108991 clone RP42-252K20 (center project name cbb)"
 25921...69950
 /note="assembly_fragment"
 70051...170904
 /note="assembly_fragment"
 150522...177535
 /note="clone overlaps with GenBank Accession Number
 AC108913 clone RP42-156J16 (center project name cbl)"
 171005...177535
 /note="assembly_fragment
 clone_end:T7
 vector_side:right"
 Query Match 56.5%; Score 2018.9; DB 1; Length 177535;
 Best Local Similarity 75.7%; Pred. No. 0;
 Matches 2572; Conservative 0; Mismatches 201; Indels 625; Gaps 12;
 QY 2 TCAGGAAGGGCGGCGAGTGAGGAGGTACTACCTCCGTCCAGGTAAGGAGCAGTAGCT 61
 Db 122254 TCAGGAAGGGCGGCGAGTGAGGAGGTAT-----CCCTGTCCCAAGGTAAGGAGCAGTGCT 122308
 QY 62 GCGCTTCTGCGAGGCGGTAAGAGATACCCCGCGCCAGGTAAAGAGAAACCAAGT 121
 Db 122309 GTGCTTTCAGGAGGAGCTGTGAAGAGATACCCCACTCCAGGTAAAGAGAAACCAAT 122368
 QY 122 AAGATGGTAGGTGTGTGAGAGGGCATCAGAGGGCAGACATACCTGAAACCATACACGAG 181
 Db 122369 AAGATGGTAGGTGTGTGCAAGAGGGCATCAGAGGGCAACACACTGAACCATATTCAG 122428
 QY 182 AAACTAGTCAATCTTAATCACTAGGACCAAGCCTTGTCTAACTTAATGAACATAGC 241
 Db 122429 AAACTAGTCAATCTTAATCACTAGGACCAAGCCTTGTCTAACTTAATGAACATAGC 122488
 QY 242 CATGCCGTGGGGCAACCAAGATGGCAGGTCTAGTGGAGAGATCTGACAGAAATGTGG 301
 Db 122489 CATGCTGTGGGGCCACCACAGACGGGCGGTGTCATGTGGAGAGATCTGACAGAAATGTGG 122548
 QY 302 TCACCTGGAGAGGGAAT-GCAAAACATTCAGTATTTCTGCTTTGAGAACCCCATGAAC 360
 Db 122549 TCACCTGGAGAGGGAATAGCAACCACTTCAGTATTTCTGCTTTGAGAACCCCATGAAC 122608
 QY 361 ACTATGAAAGGCAAAATGATAGGATCTGAAGAGGAACTCCCGAGTCTAGTAGTGCC 420
 Db 122609 AGTATGAAAGGCAAAATGATAGGATCTGAAGAGGAACTCCCGAGTCTAGTAGTGCC 122668
 QY 421 CCATATGCTTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCA 480
 Db 122669 CAATATGCTTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCA 122728
 QY 481 AAGCAAAAGAAATACCCAGCTGTGGATGTGACTGTGTATATAGCAAGGTCCGATGCTGT 540
 Db 122729 AAGCAAAAGCAATACCTAGTGTGGATGTGACTGTGTATATAGCAAGGTCCGATGCTGT 122788
 QY 541 AAGAGCAATATTGTCATAGGAACCTGGAAATGTCAAGTCCATGAATCAAGGCAAAATGGAA 600
 Db 122789 AAGAGCAATATTGTCATAGGAACCTGGAAATGTCAAGTCCATGAATCAAGGCAAAATGGAA 122848
 QY 601 GTGGTCAAAAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAGGCAACTAAAA 660
 Db 122849 GTGGTCAAAAGAGATGGCAAGAGTGAATGTTCGATTTCTAGGAATCAGCAAACTAAAA 122908
 QY 661 TGGACTGGAATGGGTGAATTTAACTCAGATCACCATTATATCTACTCTGCGGCGAGAA 720

Db	122909	TGGACTGGAATGGGTGAATTTAACTCAGATGACCAATTATATCTACTACTCGGGCAGGAA	122968
Qy	721	TCCCTCAGAGAAATGGAGTAGCCCATCATGTGTCAACAAAAGAGTCCGAAATGCACTACTT	780
Db	122969	TCCCTCAGAGAAATGGAGTAGCCCATCATGTGTCAACAAAAGAGTCCGAAATGCACTACTT	123028
Qy	781	GGATGCAGTCTCAAAAACGACAGATGATCTCTGTGTTTTCGAAAGCAACCTTCAAT	840
Db	123029	GGATGCAGTCTCAAAAACGACAGATGATCTCTGTGTTTTCGAAAGCAACCTTCAAT	123088
Qy	841	ATCAGATTAATCCCAAGTCTATGCCCCCAACCACTAGTATGCTGAAGAGCTGAAGTTGAACGG	900
Db	123089	ATCAGATTAATCCCAAGTCTATGCCCCCAACCACTAGTATGCTGAAGAGCTGAAGTTGAACGG	123148
Qy	901	TCCTATGAAGACCTTACAAGACCTTTTATGAACCTAACACCCAAAAAGATGCTTCTCAT	960
Db	123149	TCCTATGAAGACCTTACAAGACCTTTTATGAACCTAACACCCAAAAAGATGCTTCTCAT	123208
Qy	961	ATAGGGGACTGGAATGCAAAAAGTAGGAGAGCAAGAAACACCTGAGTAAACAGGCAAAATTT	1020
Db	123209	ATAGGGGACTGGAATGCAAAAAGTAGGAGAGCAAGAAACACCTGAGTAAACAGGCAAAATTT	123268
Qy	1021	GGCCTTGGAAATACGGAATGAAGAGGGCAAAAGACTAATAGAGTTTTCGCAAGAAATGCA	1080
Db	123269	GGCCTTGGAAATACGGAATGAAGAGGGCAAAAGACTAATAGAGTTTTCGCAAGAAATGCA	123328
Qy	1081	CTGTCTATAGCAACACCTCTTCCAAACACACAGAGAGACTCTACATGACATCA	1140
Db	123329	CTGTCTATAGCAACACCTCTTCCAAACACACAGAGAGACTCTACATGACATCA	123388
Qy	1141	CCAGATGCTCAACACCGAAATCAGATTGATTATATCTTTGACGCCAAAGATGAGAAAGC	1200
Db	123389	CCAGATGCTCAACACCGAAATCAGATTGATTATATCTTTGACGCCAAAGATGAGAAAGC	123448
Qy	1201	TCTATACAGTCAGCAAAAACCAAGACCGAGGCTTACTGTGGCTCAGATCATGAACTCCTT	1260
Db	123449	TCTATACAGTCAGCAAAAACCAAGACCGAGGCTTACTGTGGCTCAGATCATGAACTCCTT	123508
Qy	1261	ATTGCCAAATTCAGACTTAAATTCGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAA	1320
Db	123509	ATTGCCAAATTCAGACTTAAATTCGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAA	123568
Qy	1321	GACCTAAATCCCAATCCCTTATGATTATACAGTGAAGTAGAAATAGATTTAAGGGCCTA	1380
Db	123569	GACCTAAATCCCAATCCCTTATGATTATACAGTGAAGTAGAAATAGATTTAAGGGCCTA	123628
Qy	1381	GATCTGATAGACAGATACCTTAATGAACACTATGACAGAGTTCTATGATGTACAGGAG	1440
Db	123629	GATCTGATAGACAGATACCTTAATGAACACTATGACAGAGTTCTATGATGTACAGGAG	123688
Qy	1441	ACAGGGATCGAGACCATCCCATGGAAGAAATGTC-AAAAAGCAAAATGGCTGTCTGG	1499
Db	123689	ACAGGGATCGAGACCATCCCGTGGAAAGAAATGCAAAAGCAAAATGGCTGTCTGG	123748
Qy	1500	GGAGGCTTACAATAGCTGTGAAAGAGAGAGTGAAGAGCAAGGAAAGGAAAG	1559
Db	123749	GGAGGCTTACAATAGCTGTGAAAGAGAGAGTGAAGAGCAAGGAAAGGAAAG	123808
Qy	1560	ATAAAGCATCTGAATGCAGAGTTCCAAA	1588
Db	123809	ATAAAGCATCTGAATGCAGAGTTCCAAA	123868
Qy	1589	-----	1588
Db	123869	CAGCGATCAATGCAAGAAATAGAGGGAAGCAACAGAAATGGGAAAGACTAGAGATATCTT	123928
Qy	1589	-----	1588
Db	123929	CAAGAAATTAGAGATACCAAGGGAACATTAAACGTAAGATGGGCTCGATAAAGGCGAG	123988
Qy	1589	-----	1588
Db	123989	AAATGGTATGGACCTTAACAGAGCAGAGATATTAAGAGAGATGGCAAGATACACAGA	124048
Qy	1589	-----	1588
Db	124049	AGAACTGTACAAAAAGTTCTTCAAGACCAAGATAATCACAATGGTGTGATCACTGACCT	124108
Qy	1589	-----	1588
Db	124109	AGAGCCAGACATCTCTGGAATGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTACGAACAA	124168
Qy	1589	-----	1588
Db	124169	AGCTAGTGGAGGTGATGGAATTCAGATTCCAGCTCTTCAAAATCCTGAAAGATGATGCTGT	124228
Qy	1589	-----	1588
Db	124229	GAAAGTGTACACTCAATATGCGAGCAAAATTTGCAAAACTCAGCAGTGGCCACAGGACTG	124288
Qy	1589	-----	1588
Db	124289	GAAAACGTGAGTTTTCATTCCTCAATCCCAAGAAAGGCAATGCCAAAGAAATGCTCAAACTA	124348
Qy	1589	-----	1588
Db	124349	CTGCACAATTGTACTCATCTCATGTAGTAAAGTAAATGCTCAAAATTTCTCCAAGCCAG	124408
Qy	1589	-----	1588
Db	124409	GCTTCAGCAATATGTGAACCGTGAACTTCCAGATGTTCAAGCTGGTTTTAGAAAAAGGCGAG	124468
Qy	1589	-----	1588
Db	124469	AGAAACAGAGATCAAAATTTGCCAATCCGCTGGATCATAGAAAGCAAGAGAGTTGCA	124528
Qy	1589	-----	1588
Db	124529	GAAAAACATCTATTTCTGCTTTTATTGACTATGCCAAAGGCTT- - - - -TGTTGTTGTTTCA	124584
Qy	1589	-----	1588
Db	124585	TAAACTGTGGAATAATCTTCAAGAGATGGAAATACCAAGACCACTGACCTGACTCTTTGAA	1606
Qy	1589	-----	1588
Db	124586	AAATTTGTATGAGTGTGAGGACCAACAGTGTAGAACTGGACATGGAACCAACAGAGCTGTT	1666
Qy	1589	-----	1588
Db	124645	AAACCTGTATGAGGTGAGAAAGCAACAGTGTAGAACTGAGCATGGAACCAACAACTGTT	124704
Qy	1589	-----	1588
Db	124705	CCAAATAGGAAAGGAGTCCATCAAGGCTGTATATTGTACCCCTGCTTATTTAACTTATA	124764
Qy	1589	-----	1588
Db	124765	TGCAGAG-ACATCATGAGAAAGCTGGGCTGGAAGAAAGCAAGCTGGAATCAAGATTGC	1985
Qy	1589	-----	1588
Db	124824	TGCAGAGCAGTCAAGAGAAATGCTGGGCTGGAAGAGCAACAACTGGAATCAAGATTTC	124884
Qy	1589	-----	1588
Db	124885	CGGAGAAATAGCAATAACCTCAGATAAGCAGA-TGATACCACTTATGGCAGAAAGTG	2044
Qy	1589	-----	1588
Db	124886	TGGGAGAAATATCAATAACCTCAGATAAGCAGA-TGATACCACTTATGGCAGAAAGTG	2104
Qy	1589	-----	1588
Db	124944	AGAGGAACTAAAGGCTTCTTATGAGGTGAAGTGGAGAGTGAAGAGTTGCTTAA	2164
Qy	1589	-----	1588
Db	124945	AGCTCAACATTCAGAAACGAAAGATCATGGCTGCTGCTCCATCATCTTATGGGAAATAG	125004
Qy	1589	-----	125004
Db	125005	ATGGGAAACAGTGGAAATAGTGTGAGAGTTT- - - - -TTTTGGGGGGCTCCAAATC	2224
Qy	1589	-----	125057
Db	125058	AGATGCTGACTGACGCAATGAAATTAAGAGACACTTACTCTTTGGAAGAAAGTTA--AC	2282
Qy	1589	-----	125058
Db	125059	ACTGAGTCAATGAAATTAAGAGATGCT-TACTCTTTTGGAGAAAGTTATGAC	125108
Qy	1589	-----	125108

Db	166527	CCAGGTAGTAGGGCCCAATATGCTACTGAGATCAGTGGAGAAATAACTCCAGAAGA	166586	Db	167605	GCAAAATGGCTGTATGGGGAGGCTTTACAAATAGCTGTGAAAAAGAGAGTGAAGC	167664
Qy	464	ATGAAGAGATGGAGCCAAAGCAAAAGAAATACCCAGCTGTGGATGTGCTGGTATATA	523	Qy	1543	AAAGGAAAAAGAGAAAGATTAAGCATCTGAATGCGAGATTCCAAA	1588
Db	166587	ATGAAGGATGGAACCAAGCAAAACCAATACCCAGCTGTGGATGTGCTGGTATAGAA	166646	Db	167665	AAAGGGCAAAAGGAAAGATGTAAGCATCTGAATGCGAGATTCCAAAAGATAGCCAGAAGA	167724
Qy	524	GCAAGTCCGATGCTGTATGAAGAGCAATATTGCAAGGAACCTGGAATGTCAAGTCCCATGA	583	Qy	1589	-----	1588
Db	166647	GCAAGTCCCATGCTGTGAAGGCAATGTTTCATAGAAACCTGGAAATGTCAGGCCCATGA	166706	Db	167725	GATAAGAAAGCCTTTCTCATCGATCAGTGCMAAGAAATAGAGGTAAACAACATGAATGGGA	167784
Qy	584	ATCAAGGCAAAATGGAAGTGTCAAAAGAGATGGCAAGAGTGAATGTCAACATTTCTAG	643	Qy	1589	-----	1588
Db	166707	ATCAAGGCAAAATGGAAGTGTCAAAAGAGATGGCAAGAGTGAATGTCAACATTTCTAG	166766	Db	167785	AAGACTAGAGATCTTTCAAGAAATCAGAGATACCAAGGGGACATTTTCATGCAAGATG	167844
Qy	644	GAATCAGCGAACTAAATGGAGCTGGAATGGGTGAATTTAACTCAGATGACCATATATCT	703	Qy	1589	-----	1588
Db	166767	GAATCAGCGAACTGAATGGAGTGAATGGGTGAATTTAACTCAGATGACCATATATCT	166826	Db	167845	GGCTCGATAAGGACAGAAATGGTATGGATCTAAACAGAGCAGAGATATTAAAGGAGGTG	167904
Qy	704	ACTACTCGGGCAGGAATCCCTCAGAGAAATGGAAGTAGCCATCATGTCAACAAAAGAG	763	Qy	1589	-----	1588
Db	166827	ACTACTCGGGCAGGAATCCCTCAGAGAAATGGAAGTAGCCATCATGTCAACAAAAGAG	166886	Db	167905	GCAAGAATACACAGAGAACTGTACAAAAGATCTTCAGACCAGATAATCAGCAGCG	167964
Qy	764	TCCGAAATGCAGTACTTGGATGCAATCTCAAAAAGCAGAGATGATCTCTGTTTCTCC	823	Qy	1589	-----	1588
Db	166887	TCCGAAATGCAGTACTTGGATGCAATCTCAAAAAGCAGAGATGATCTCTGTTTCTCC	166946	Db	167965	TGTGATCACTGACCTAGGTTGACATCTCTGGAATGTGAAGTCAAGTGGGCTTAGAAAG	168024
Qy	824	AAGCAAAACCATTAATACAGTAATCCAAAGTCTATGCCCCCAACCAAGTAATGCTGAAG	883	Qy	1589	-----	1588
Db	166947	AAGCAAAACCATTAATACAGTAATCCAAAGTCTATGCCCCCAACCAAGTAATGCTGAAG	167006	Db	168025	CATCACTACGAAACAAAGCTAGTGGAGGTGATGGAATTCAGTCGAGCTATTTCAAATCCT	168084
Qy	884	AAGCTGAAGTTGAACGGTCTTATGAAGACCTTAAGACCTTTTGAATTAACACCCAAAA	943	Qy	1589	-----	1588
Db	167007	AAGCTGAAGTTGAATGGTCTTATGAAGACCTTACAAAGACCTTTTGAATTAACACCCAAAA	167066	Db	168085	GAAAGATGATGTGTTAAAAAGTCTGCATCAATATGCCAGCAAAATTTGAAAACTCAGCA	168144
Qy	944	AAGATGCTCTTCTCATATAGGGACTGGAATGCAAAAGTAGGAAGCAAAACACCTG	1003	Qy	1589	-----	1588
Db	167067	AAGATGCTCTTCTCATATAGGGACTGGAATGCAAAAGTAGGAAGCAAAACACCTG	167126	Db	168145	GTGCCCACAGATTGGAAGGTCAGTTTTCATTCCAATCCCAAGGAGCAATGCTCAA	168204
Qy	1004	GAGTAACAGGCAAAATTTGGCTTGGATACGGAATGAAGAGGGCAAGCAATATAGT	1063	Qy	1589	-----	1588
Db	167127	GAGTAACAGGCAAAATTTGGCTTGGATACGGAATGAAGAGGGCAAGCAATATAGT	167186	Db	168205	AGAATCCTCAAACTACCGCACAAATTCGACTCATCTCATGTGTAGTAAAGTAATGCTCAA	168264
Qy	1064	TTTGCCAAAGAAATGCACTGCTATAGCAACACCTCTTCCAAACACCAAGAGAGAC	1123	Qy	1589	-----	1588
Db	167187	TTTGCCAAAGAAATGCACTGCTATAGCAACACCTCTTCCAAACACCAAGAGAGAC	167244	Db	168265	AAATTCCTCAAGCCAGCTTCAGCAATACGTGAACCGTGAATTCAGATGTTAAAGCTGG	168324
Qy	1124	TCTACATGGACATCACCAGATGTCACACCGAAATCAGATTATATTTTGGCA	1183	Qy	1589	-----	1588
Db	167245	TCTACATGGACATCACCAGATGTCACACCGAAATCAGATTATATTTTGGCA	167304	Db	168325	TTTTAGAAAAGGTCAGAGAACCAAGATATCAAAATTCGCAACATCCGCTGGATCATGGAAA	168384
Qy	1184	GCCAAAGATGGAGAGCTCTATACAGTCAGCAAAACCAAGACAGGAGCTTACTGTGGCT	1243	Qy	1672	AGCAAGAGATTCACAGAAAACATCTATTTCTGCTTTATTGACTATGCAAAAGCCTTGA	1731
Db	167305	GCCAAAGATGGAGAGCTCTATACAGTCAGCAAAACCAAGACAGGAGCTTACTGTGGCT	167364	Db	168385	AGCAAGGAGTTCCAGAAAACATCTATTTTGTCTTTATTTGACTATACCAAGCCTTTGA	168444
Qy	1244	CAGATCATGAATCTCTTATGCAAAATTCAGACTTAAATTCAGAAAGTAGGGAAAAACCA	1303	Qy	1732	CTGTGGGGTTCACAAATACTGGAAAATCTGAAAAGGATGGGAATACCAAGACCACT	1791
Db	167365	CAGATCATGAATCTCTTATACCAATTCAGACCTTAATTTGAAGAGTAGGGAAAAACCA	167424	Db	168445	CTGTGGATCACAATAAAC--TGGACATTTCTGAAAAGTGGGNAATACCAAGACCACT	168502
Qy	1304	CTAGATCTCAGGTAGACCTTAATCCAAATCCCTTATGATTAATACAGTGAAGTGAAGAA	1363	Qy	1792	GACTGACTCTTGAAAAATTTGTATGCAAGTCAGAGAACCAAGTTAGAACTGGACATGG	1851
Db	167425	CTAGACCATTCAGGTATGACCTTAATCCAAATCCCTTATGATTAATACAGTGAAGTGAAGAA	167484	Db	168503	GATCTGCCTCTTGAGAAATCTGTATGCAAGTCAGAGAACCAAGTTAGAACTGGACATGG	168562
Qy	1364	ATAGATTTAAGGCTTAGATCTGTAGAG--CAGAGTACCTTAATGAATGACATGACAGAGGTT	1422	Qy	1852	AACAACAGACTGTTCCCAAGTAGGAAAAGAGTATGTCAAGGCTGTATATTTGTCAACCCGG	1911
Db	167485	TTAGATTTAAGGCTTAGATCTGTAGAGAGAGAGTGGCTGTATGAATGGAATGAGGTT	167544	Db	168563	AACAACAGACTGTTCCCAATAGGAAAAGAGTATGTCAAGGCTGTATATTTATTCACCTGT	168622
Qy	1423	CATGACATTTGACAGGACAGGATCGAGACCACTCCCATGGAAAAAGAAATGCAAAAAA	1482	Qy	1912	CTTGTTTAACTTATGTCAGAG--ACATCATGAGAAACCGCTGGGCTGGAAAGAGCAACAG	1970
Db	167545	CGTGACATTTGACAGGACAGGATCAAGACCACTCCCATGGAAAAAGAAATGCAAAAAA	167604	Db	168623	TTTATTTAACTTATATGACAGGTACATCATGAGAAATGCTGGACTGGAGAAACACAAAC	168682
Qy	1483	GCRAATGGCTGTCTGGGGAGCCCTTACAAATAGCTGTGAAAAAGAGAGTGAAGC	1542	Qy	1971	TGGAATCAAGATTCGCGGAGAAATAGCAATTAACCTCAGATATGCAATATACCACTCT	2030
				Db	168683	TGGAATCAAGATTCGCGGAGAAATATCAATTAACCTCAGATATGCAATATACCACTCT	168742

2031 TATGCGAAGAGTGAAGAGGAACTAAAGGCTCTTGATGAAGTGAAGAGGAGAGTGA 2090
Db 168743 TATGCGAAGAGTGAAGAGGAA-TAAAGGCTCTTGATGAAGTGAAGAGGAGTGA 168801
Qy 2091 AAAAGTTGGCTTAAAGCTCAACATTCAGAAAACGAGATCATGCATCTGCTCCATCAC 2150
Db 168802 AAAAGTTGGCTTAAAGCTCAACATTCAGAAAACGAGATCATGCATCTGCTCCATCAC 168861
Qy 2151 TTATGCGAAGTGAAGAGGAAACAGTGGAAACAGTGCAGATTTATTTTGGGGGGC 2210
Db 168862 TTATGCGAAGTGAAGAGGAAACAGTGGAAACAGTGCAGATTTATTTTGGGGGGC 168919
Qy 2211 TCCAAATTCAGTGCAGATGCTGACGACCATGAATTAAGACACTTACTCTCTTGGGA 2270
Db 168920 TCCAAATTCAGTGCAGATGCTGACGACCATGAATTAAGACACTTACTCTCTTGGGA 168979
Qy 2271 AGAAAGTTA--ACCAACTAGATAGCATATTGAAAGCAGAGACATTTACCTCCCAACA 2328
Db 168980 AGAAAGTTATGACAAACCTAGATAGCATATTGAAAGCAGAGACATTTACCTCCCAACA 169039
Qy 2329 AAGCCCACTAGTCAAGCTATGTTTTTTCAGTGTCTATGATGTGAGATGG 2388
Db 169040 AAGTCCCTCTAGTCAAGCTATGTTTTTTCAGTGTCTATGATGTGAGATGG 169099
Qy 2389 ACTGTGAAGAAAGCTGACACTGAAGAATTGATGCTTTTGAACCTGTGTGTGGAGAAGA 2448
Db 169100 ACTGTGAAGAAAGCTGACGCGCAAGATTTGCTTTTGAACCTGTGTGTGGAGAAGA 169159
Qy 2449 CTCCTGAGAGTCCCTTGGACTGCAAGAGATGCAACGATCCATTCGAGGATGATCAGC 2508
Db 169160 CTCCTGAGAGTCCCTTGGACTGCAAGAGATGCAACGATCCATTCGAGGATGATCAGC 169219
Qy 2509 CTTGGGATTTCTTGGAAAGGATGATGCTAAAGCTGAACCTCCAGTACTTTGGCCACTG 2568
Db 169220 CTTGGGATTTCTTGGAAAGGATGATGCTAAAGCTGAACCTCCAGTACTTTGGCCACTG 169279
Qy 2569 ATCAGAAGAGTGACTCACTCGAAAGACCTGATGCTGGAGGATTTGGGGCAGGAGG 2628
Db 169280 ATGTCAGAGAGTGACTCACTCGAAAGACCTGATGCTGGAGGATTTGGGGCAGGAGG 169339
Qy 2629 AGAAGGAGCAGCAGAGATGATGCTGGATGCGATCACTGATGATGATGATGATGATGAT 2688
Db 169340 AGAAGGAGCAGCAGAGATGATGCTGGATGCGATCACTGATGATGATGATGATGATGAT 169398
Qy 2689 CTGGGTGAACCTCTCGAGTGTGTGATGACAGGAGGCTGCTCTCGCGGATTCATGGG 2748
Db 169399 CTGAGTGAACCTCGGAGGATTTGTGATATACAGGAGGCTGCTCTCGCGGATTCATGGG 169458
Qy 2749 GTCAAAAGAGTTGACAGCAGTGAAGAACTGAACCTGAACCTGAACCTGAACCTGA 2795
Db 169459 GTCGAAAGAGTTGACAGCAGCAGCTAGTGAACCTGAACCTGAACCTGAACCTGA 169505

RESULT 30
AC137651
LOCUS Bos taurus clone RP42-22413, WORKING DRAFT SEQUENCE, 2 ordered pieces.
DEFINITION AC137651
ACCESSION AC137651
VERSION AC137651.2 GI:28394356
KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 192540)
Benjamin, B., Blakesley, R.W., Bouffard, G., Brinkley, C., Brooks, S.,
Carliaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Margulies, E.H., Masiello, C., Maskeri, B., McDowell, J.,
Paguirigan, C., Pearson, R., Portnov, M.E., Prasad, A.,
Reddix-Dugue, N., Schandler, K., Schueler, X.G., Sison, C.,
Stankis, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 192540)
Green, E.D.
Direct Submission
Submitted (27-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 192540)
Green, E.D.
Direct Submission
Submitted (15-FEB-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Feb 15, 2003 this sequence version replaced gi:25699528.
----- Genome Center
Center: NIH Intramural Sequencing Center
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: crw
Center clone name: 224103

The sequence data in this record represents an 'enhanced' version of a phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; r/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 192045 bases at least Q40
Consensus quality: 192329 bases at least Q30
Consensus quality: 192389 bases at least Q20
Insert size: 167000; agarose-fp
Quality coverage: 12.00x in Q20 bases; agarose-fp
Quality coverage: 10.41x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.
* 1 145285: contig of 145285 bp in length
* 145286 145385: gap of unknown length
* 145386 192540: contig of 47155 bp in length.

FEATURES
source
1..192540
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-22413"
/clone_lib="RP42"
1..145285
misc_feature
/note="assembly_fragment
clone_end:SP6

misc_feature		vector_side:left"		vector_side:right"	
145386..192540		54.5%; Score 1945.1; DB 1; Length 192540;			
/note="assembly_fragment		74.9%; Pred. No. 0;			
clone_end:17		0; Mismatches 214; Indels 617; Gaps 10;			

1589 -----GAACTTCC 1596
19390 TAAATAATGCTCAAAATTTCTCAAGCCAGGCTTTCAGCAATACGTTGAACCATGAAATTC 19449
1597 AGTTGTTCAAGCTGGTTTGAAGAAGTCAGAGGAACAGAGACCAAAATTTGCCACATCCT 1656
19450 AGATGTTCAAGCCAGTTTTCGAAGAAGCAGAGAACCAAGATCAAAATTTGCCACATTTG 19509
1657 CTGTATCATGGAAGAACAGAGAGTTCCAGAAAAACATCTATTCTGCTTTATTGACTA 1716
19510 CTGGATCATCAAAAAGCAAGAGAGTTCCAAAAAACATCTATTCTGCTTTATTGACTA 19569
1717 TGCAAAAGCCTTTGACTGTGGGGTCAAAATAAATCTGGAATTTCTGAAAGGATGGG 1776
19570 TGCCAAAGCCTTTGACTGTGGATCAAAATAAATCTGGAATTTCTGAAAGGATGGG 19629
1777 AATACCAAGCACCCTGACTGCTTCAAAAATTTGTATGAGGTTCAGGAAGCAACAGT 1836
19630 AATATCAGACCACTGACTTGCCTCTTGAGAACCTATATGCAGGTTCAGGAAGCAACAGT 19689
1837 TAGAATCGACATGGAACACAGAGCTGTTCAAGTAGGAAGAACAGATATGTCAGGCTG 1896
19690 TAGAATCGACATGGAACACAGAGCTGTTCCAAATGGAAGAACAGATATGTCAGGCTG 19749
1897 TATATTGTCACCGCTTGTGTTTAACTCTATGCGAG-ACATCATGAGAAACGCTGGCT 1955
19750 TATATTGTCACCGCTTGTGTTTAACTCTATGCGAGTACATCATGAGAAACGCTGGCT 19809
1956 GGAAGAAGCACAAGCTGGAATCAAGATTGCCGGGAGAAATAGCAATTAACCTCAGATATGC 2015
19810 AGAGGAAGCACAAGCTGGAATCAAGATTGCCGGGAGAAATCTCAATTAACCTCAGATATGC 19869
2016 AGATATACCACTTATGCGGAAGTTCAGAGGAACCTTAAGAGCTTTCAGAGGT 2075
19870 AGATATACCACTTATGCGGAAGTTCAGAGGAACCTTAAGAGTCTCTTTCAGAGGT 19829
2076 GAAAGAGGAGAGTGAAGAGTTGGCTTAAAGCTCAACATTCAGAAAAAGCAAGATCATGGC 2135
19930 GAAAGAGGAGAGTGAAGAGTTGGCTTAAAGCTCAACATTCAGAAAAAGCAAGATCATGGC 19989
2136 ATCTGCTCCATCATCTTCATGGAATATAGATGGGAAACAGTGGAAACAGTTCAGACTT 2195
19990 ATCTAGTCCATCATCTTCATGGAATATAGATGGGAAACAGTGGAAACAGTTCAGACTT 20049
2196 TATTTTGGGGGGTCCAAATCACTGAGATGCTGAGATGCTGAGCAATGAAATTAAGA 2255
20050 TA--TTTTGGGGGGTCCAAATCACTGAGATGCTGAGTGCAGCCATG----- 20096
2256 CACTTACTCTTTGGAAGAAAGTTA--ACCAACCTAGATAGATATGGAAGAGCAGAGAC 2313
20097 -GTTTACTCTTTGGAAGAAATATATGACCAAACTAGATAGATATGGAAGAGCAGAGAC 20155
2314 ATTACTTTGCCAAAGCCCATCTAGTCAGGCTATGTTTTCAGTGGTCTATGAT 2373
20156 ATTACTTTGCCAAAGCCCATCTAGTCAGGCTATGATTTTTCAGTGGTCTATGAT 20215
2374 GGATGTGAGATGTGACTGTGAAGAAAGCTGAGCACTGAAGAAATGATGCTTTTGAATG 2433
20216 GGATGTGAGATGTGACTGTGAAGAAAGCTGAGCACTGAAGAAATGATGCTTTTGAATG 20275
2434 TGGTGTGAGAGACTCTTGAGTCCCTTGGCTGCAAGAGATCCAAACGATTCATT 2493
20276 TGGTGTGAGAGACTCTTTAGATCCCTTGGCTGCAAGAGATCCAAACGATTCATT 20335
2494 CTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATCTTAAAGCTGAAATCCAG 2553
20336 CTAAGAGAGATCAGTCTCGGTGTTCAATGGAAGAAC-TATGCTGAGTGAATCCAA 20395
2554 TACTTTGGCCACTGATTCAGAGAGCTGATCACTGGAAGAACCTGATGCTGGAGGG 2613
20396 TACTTTGGCCACTGATTCAGAGAGCTGATCACTTGGAAAGAACCTGATGCTGGAGGG 20455
2614 ATTGGGGGCGAGGAGAGGGGACGACAGAGGATGAGATGGCTGGATGGCATCTACGAC 2673

20456 ATTGGGGCGAGAGAGAGAGGAGGAGAGAGATGAGATGGCTGATGGCATCACCGAC 20515
2674 TCGATGAGCGTGAAGTCTGGTGAATCTCTGGAGTTGGTGAAGAGAGGAGGCGCTGTCT 2733
20516 TCAATGAGCATGAGTTTGAATTAATCTCTGGAGTTTGAATTAATGAACAGGAGGCGCTGGCGT 20575
2734 GCGGCAATTCATGGGTCACAAAGAGTTGGACACGACTGAGCAACTGAACCTGAACCTGAAC 2793
20576 GCTGTGAATTCCTGGGTCACAAAGAGTCAGACACGACTGAGCAACTGAACCTGAACCTGAAC 20635
2794 TGTACTGAAACCT 2806
20636 TGAACCTGAACCAT 20648
AC092727 163604 bp DNA linear MAM 10-OCT-2002
Bos taurus clone RP42-351K5, complete sequence.
AC092727
AC092727.2 GI:17386265
HTG.
Bos taurus (cow)
Bos taurus
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 163604)
Akher, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granitz, S., Guan, X., Gupta, J., Haghighi, P.,
Ho, S.-L., Idol, J.R., Karlins, E., Laric, P., Lee-Lin, S.-Q.,
Legaspi, R., Maduro, Q.L., Maduro, V.B., Masello, C., Mastrian, S.D.,
McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A., Shevchenko, Y.,
Stancic, S., Thomas, J.W., Thomas, P.J., Touchman, J.W.,
Tsaur, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 163604)
Green, E.D.
Direct Submission
Submitted (21-JUL-2001) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 163604)
Green, E.D.
Direct Submission
Submitted (06-DEC-2001) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
4 (bases 1 to 163604)
Green, E.D.
Direct Submission
Submitted (10-OCT-2001) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
On Dec 6, 2001 this sequence version replaced gi:14993717.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
----- Project Information
Center project name: cjb
Center clone name: 351K05

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the features section.

FEATURES

Query Match	Best Local Similarity	Score	DB 1;	Length
Matches 2468;	Conservative	0;	Mismatches 153;	Indels 629;
Gaps				
162	TACTGAAACCATACACGACGAGAAAACCTAGTCAATCTAATCACACTAGGACCCACAGCCTTGT	221		
42552	TAGTGGCCAGTCACAGCCNAAACTAGTCAATCTAATCACACTAGGACCCACAGCCTTGT	42493		
222	CTAACTCAATGAACTAAGCCATGCCCCGTGGGGCAACCCAAAGATGGGAGGTCATGGTGG	281		
42492	CTAACTCAATGAACTAAGCCATGCCCCGTGGGGCAACCCAAAGATGGGAGGTCATGGTGG	42433		
282	AGAGATCTGACAAATGTGTCTCCACCTGGAGAGAGGAAT-GCAAAACACTTCAGTATCTCT	340		
42432	AGAGGTCTGACAAATGTGTCTCCACCTGGAGAGAGGAATGCGAAACCACTTCAGTATCTCT	42373		
341	GCCTTGAGAACCCCATGAACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAAC	400		
42372	GCCTTGAGAACCCCATGAACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAAC	42313		
401	TCCCGAGTCAGTAGGTGCCCATATGCTACTGGAGATCAGTGGAGAAATTAATCTCCAGAA	460		
42312	TCCCGAGTCAGTAGGTGCCCATATGCTACTGGAGATCAGTGGAGAAATTAATCTCCAGAA	42253		
461	AGAAATGAAGAGATGGAGCCCAAGCAAAAGAAATATCCAGCTGTGGATGTGACTGTGATA	520		
42252	AGAAATGAAGAGATGGAGCCCAAGCAAAAGAAATATCCAGCTGTGGATGTGACTGTGATA	42193		
521	TGACGAGGTCCGATGCTGTATGAGACCAATATGCTATGGACACCTGGAGATGTCAGGTCCA	580		
42192	GAAGCAAGGTCCGATGCTGTATGAGACCAATATGCTATGGACACCTGGAGATGTCAGGTCCA	42133		
581	TGAATCAAGGCAAAATGGAAAGTGGTCAACCAAGAGATGGCAAGAGTGAATGTCAACATTC	640		
42132	TGAATCAAGGCAAAATGGAAAGTGGTCAACCAAGAGATGGCAAGAGTGAATGTCAACATTC	42073		
641	TAGGAATCAGGAACTTAAATGAGCTGGATGGTGAATTTAACTCAGATGACCAATTATA	700		
42072	TAGGAATCAGGAACTTAAATGAGCTGGATGGTGAATTTAACTCAGATGACCAATTATA	42013		
701	TCTACTACTGCGGCGAGGAATCCCTCAGAGAGAAATGGA--GTAGGCATCATGCTCAACAA	758		

42012 TCTACTACTGGGGCAGGAATCCCTCAGAAAGAAACGGACTGTAGCCATCGTGGTCAACAA 41995

759 AAGAGTCCGAAATGCACTACTTGGATGCGACTCTCAAAAACGACAGAAATGATCTCTGTTG 818

41952 AAGAGTCTGAAATGCACTACTTGGATGCGAAATCTCAAAAACGACAGAAATGATCTCTGTTG 41893

819 TTTCCAGGCGAAACCAATTCAAATACACAGTATCCAAAGTCTATGCCCAACAGTAATGC 878

41892 TTGCGAAGGCGAAACCAATTCAAATACACAGTATCCAAAGTCTATGCCCAACAGTAGCAC 41833

879 TGAAGAAGCTGAAGTTGAACGGTCCCTATGAAGACCTTACAAGACCTTTTAGAACTAACACC 938

41832 TGAAGAAGCTGAAGTTGAATGGTCTTATGAAGACCTTACAAGACCTTTTAGAACTAACACC 41773

939 CAAAAAAGATGTCTCTTCATATAGGGACCTGGAAATGCAAAAGTAGAGAGCAAGAAAC 998

41772 CAAAAAAGACGTCTCTTTTCATATAGGGACCTGGAAATGCAAAAGTAGAGAGTCAAGAAC 47113

999 ACCTGGAGTAAACAGGCAAAATTTGGCTTTGGAAATACGGAATGAAGCAGGGCAAGACTAAT 10558

41712 ACCTGGAGTAAACAGGCAAAATTTGGCTTTGGAATACGGAATGAAGCAGGGCAAGACTAAT 41653

1059 AGAGTTTTCGCAAGAAATGCACTGGTCTATAGCAAAACACCCCTCTTCCAAACAACAAGAG 1118

41652 AGAGTTTTCGCAAGAAATGCACTGGTCTATAGCAAAATGCACTCTTCCAAACAACAAGAG 41593

1119 AAGACTCTACACATGGACATCACAGATGCTCAACACCGAAATTCAGATTGATTATTTCT 1178

41592 AAGA - CTACACATGGACATCACAGATGCTCAACACCGAAATTCAGATTGATTATTTCT 41535

1179 TTGCAGCCAAAGATGGAGAAGCTCTATACAGTCAGCAAAAACAAGCCAGGAGCTTACTG 1238

41534 TTGCAGCCAAAGATGGAGAAGCTCTATACAGTCAGCAAAAACAAGCCAGGAGCTTACTG 41475

1239 TGCTCAGATCATGAACTCTCTTATGGCAATTCAGACTTAAATTCAGAGAAAGTAGGAA 1298

41474 TGCTCAGATCATGAACTCTCTTATACCAAAATTCAGACTTAAATTCAGAGAAAGTAGGAA 41415

1299 AACCACTAGATCACTCAGSTAAGAACCCTTAAATCCAAATCCCTTATGATTATACAGTGAAGT 1358

41414 AACCACTAGATCACTCAGSTAAGAACCCTTAAATCCAAATCCCTTATGATTATACAGTGAAGT 41355

1359 GAGAAATAGATTTAAGGCCCTAGATCTGATAGACAGAGTACCTAATGAATCTATGNCAGA 1418

41354 GAGAAATAGATTTAAG - CCTAGATCTGATAGATAGAGTGCCTGATGAATCTATGGAATGA 41296

1419 GCTTCATGACATCTGTACAGGAGCAGGGATCGAGACCATCCCTGCAAAAGAAATGCAA 1478

41295 GCTTCGTCGCAATGTACAAGACAGGGATCAGACCATCCCTGCAAAAGAAATGCGA 41236

1479 AAAAGCAAAATGGCTGTCTGGGAGGCCCTTACAAATAGCTGTGAAAAGAGAAAGTGAA 1538

41235 AAAAGCAAAATGGCTGTCTGGGAGGCCCTTACAAATAGCTGTGAA - - - AAGAGAAAGCAA 41179

1539 AAGCAAAAGGAAAAGGAAAGATAAAGCATCTGAATCGAGAGTCCAAA - - - - - 1588

41178 AAGCAAAAGGAAAAGGAAAGATAAAGCATCTGAATCGAGAGTCCAAAAGATGCAAG 41119

1589 - - - - - 1588

41118 AAGAGATAAGAAAGCCTTCTTCAGCTATCAATGCAAAAGACTAGAGGAAAACAACAGAAAT 41059

1589 - - - - - 1588

41058 GGGAAAGACTAGAGATCTCTTTCAAGAAAATTAGAGATACCAAGGGAACTTTTCATGCAA 40999

1589 - - - - - 1588

40998 GATGGCTCGATAAAGACGAAATGCTATGGACCTTAACAGAGCAGAGATATTAGAA 40939

1589 - - - - - 1588

40938 GAGGTGGCAAGATACACCGAAGAACTGTGTCAAAAAAAGATCTTTCAGATCCAGATAATCA 40879

QY 1589 ----- 1588
 Db 40878 TGATGGTGTGATCACTGACCTAGAGCCAGACATCCTGGAATGTAAAGTCAAGTGCAGCTT 40819
 QY 1589 ----- 1588
 Db 40818 AGAAGCATCACTATGACAAAGCTAGTGGAGGTGATGGAATCCAGTGGAGCTATTTC 40759
 QY 1589 ----- 1588
 Db 40758 AATCTGAAGATGATGCTGTGAAAGTGTCTGACCTCAATATGCCAGCAAAATTTGGAAGC 40699
 QY 1589 ----- 1588
 Db 40698 TCAGCAGTGGCCACAGACTAGAAAGGTGAGTTTTCATCTGATCCCAAGAAAGGCAA 40639
 QY 1589 ----- 1588
 Db 40638 TGCCAAAGATGCTCAAACTACACACAGTTGCCACCATCTCACAGCTAGTAAAGTAAT 40579
 QY 1589 ----- 1588
 Db 40578 GCTCAAAATCTCAAGCCAGCTTCAGCAATATGTGAATTTGAATCTGCTGATGTTCAA 40519
 QY 1607 GCTGGTTTTAGAAAAGTCCAGAGGAACACAGAGACCAAAATTTGCCAATCTCTGTATCATG 1666
 Db 40518 GCTGGTTTTAGAAAAGCAGAGGAACACAGAGATCAATTTGCCAATCTCTGTATCATG 40459
 QY 1667 GAAAAAGCAGAGAGTTCCAGAAAAACATCTATTTCTGCTTTATGACTATGCAAAAGCC 1726
 Db 40458 GAAAAAGCAGAGAGTTCCAGAAAAACATCTATTTCTGCTTTATGACTATGCAAAAGCC 40399
 QY 1727 TTGACTGTGGGGTCAAAATAAATCTGGAAGATTTCTGAAAGGATGGCAATACACAGAC 1786
 Db 40398 TTTGACTGTGGATCAACATTAACCTGGAATTTCTCAAGATGGGAATACACAGAC 40339
 QY 1787 CACCTGACCTGACTCTTGAAAAATTTGTATGACAGGTGAGGAAGCAACAGTTAGAACTGGA 1846
 Db 40338 CACCTGACCTGCTCTTGAAAAATTTGTATGACAGGTGAGGAAGCAACAGTTAGAACTGGA 40279
 QY 1847 CATGMAACACAGACTGTTCCAAAGTAGGAAGGATGTGCAAGGCTCTATATTTGCTCA 1906
 Db 40278 CATGMAACACAGACTGTTCCAAAGTAGGAAGGATGTGCAAGGCTCTATATTTGCTCA 40219
 QY 1907 CCCGGCTTGTTAACTCTATGACAGAG-ACATCATGAGAAACGCTGGGCTGGAAGAGCA 1965
 Db 40218 CCTGCTTATTTAACTCTATGACAGAGTACATCATGAGAAATGCTGACTAGAGAAACA 40159
 QY 1966 CAAGCTGGATCAAGATTCGGGAGAAATAGCAATTAACCTCAGATATGAGATGATACC 2025
 Db 40158 CAAGCTGGAATCAAGATTCGGGAGAAATATCAATTAACCTCAGATATGAGATGATACC 40099
 QY 2026 ACCCTTATGGCAGAAAGTGAAGAGCACTAAAGACCTCTTGATGAAGGTGGAAGAGGAG 2085
 Db 40098 ACCCTTATGGCAGAAAGTGAAGAGCACTAAAGACCTCTTGATGAAGGTGAAGTGGAG 40039
 QY 2086 AGTGAAGAAAGTGGCTTAAGCTCAACATTCAGAAAGCAAGATC-----ATGGCATCTGG 2141
 Db 40038 AGTGAAGAAATTTGGCTTAAAGCTCAACATTCAGAAAGCAAGATCATTGGATGGCATCCAG 39979
 QY 2142 TCCCATCACTTCATGGGAATAGATGGGAAACAGTGGAAACAGTGCAGACTTTATTTT 2201
 Db 39978 TCCCATCACTTCATGGGAATAG-----GGAAACAGTGCAGAAATTTATTTT 39932
 QY 2202 TTGGGGGGCTCCAAAATCACTGCAGATGGTGAATGACGCAATGAAATTAAGAAAGCACTTA 2261
 Db 39931 TT--GGGTCTCCAAAATCACTGCAGATGGTGAATGACGCAATGAAATTAAGAAAGCACTTA 39878
 QY 2262 CTCCTTCGAGAGAAAGTTA--ACCAACCTAGATAGCATATTTGAAAGCAGAGACATTAACC 2319
 Db 39877 CTCCTTCGAGAGCAAGTTTATGTTCCACGTAGATAGCATATTTGAAAGCAGAGACATTAACC 39818

QY 2320 TTGCCAACAAAGCCCATCTAGCTCAAGGCTATGTTTTTCCAGTGGTCACTGATGATGT 2379
 Db 39817 TTGCCAACAAAGGTCGCTTAGTCAAGGCTATGGTTTTTCCAGTAGTCACTGATGATGT 39758
 QY 2380 GAGAGTTGAGCTGTGAAGAAAGCTGAGCAGCTGAAGAAATGTGTTTTGAACCTGTGTGT 2439
 Db 39757 GAGAATTGGAGTGTGAAGAAAGCTGAGTGTCTGAAGAAATGTGTTTTGAACTGTGTGT 39698
 QY 2440 TGGAGAGACTCTTGAGAGTCCCTTGGACTGCAAGAGATCCAAACAGTCCATTTCTGAAG 2499
 Db 39697 TGGAGAGACTCTTGAGAGTCCCTTGGACTGCAAGAGATCCAAACAGTCCATTTCTGAAG 39638
 QY 2500 GAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTT 2559
 Db 39637 GAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTT 39578
 QY 2560 GGCCACCTGATCAGAAAGCTGACTCACTGGAAGAAAGCCCTGATGCTGGAGGATTTGGG 2619
 Db 39577 GGCTCTCTCATCGAAGAGTTGACTCATTTGGAAGAAAGCTGATGCTGGAGGATTTGGG 39518
 QY 2620 GGCAGAGAGAGAGGAGGACGACAGAGGATGAGATGCTGGATGGCATCACTGACTCGATG 2679
 Db 39517 GGCAGAGAGAGGAGGAGGACGACAGAGGATGAGATGCTGGATGGCATCACTGACTCGATG 39458
 QY 2680 GACGTGAGTCTGGGTGAACTCCTGGAGTTGTTGATGGAAGAGCCCTGATGCTGGAGGATTTGGG 2739
 Db 39457 GATGTGAGTCTGAGTGAATCTCGGAGTTGTTGATGGAAGAGCCCTGTTGTTGCTGCG 39398
 QY 2740 ATTCTATGGGTCACAAAGAGTTGGACAGACTGAG--CAACTGAACTGAACTGAACTGTA 2797
 Db 39397 ATTCTATGGGTTGCAAGAAATCGGACAGTTGAGCTCACTGAACTGAACTTGAGCAAT 39338
 QY 2798 CTGAAACCTT 2807
 Db 39337 CTTTGGCAAT 39328

RESULT 32

AC092410

LOCUS

DEFINITION

AC092410.1

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC092410 168497 bp DNA linear HTG 04-JUL-2001
 Bos taurus clone RP42-245C6, WORKING DRAFT SEQUENCE, 5 unordered pieces.
 AC092410.1 GI:14595784
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 Bos taurus (cow)
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
 1 (bases 1 to 168497)
 Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Lin, S.-Q., Legaspi, R., Lim, M., Maduro, O.L., Maduro, V.B., Masello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A., Shevchenko, Y., Snyder, B., Stantropop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsurgan, L.H., Vogt, J.L., Walker, M.A., Wetherby, K.D., Zhang, L.-H. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 168497)
 Green, E.D.
 Direct Submission
 Submitted (04-JUL-2001) NIH Intramural Sequencing Center, 8717 Grovmont Circle, Gaithersburg, MD 20877, USA
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.mouse@nihri.nih.gov
 ----- Project Information
 Center project name: cjc

Db 144818 CAAGGAGAGATGAAGAAAGCCTTCTCAGCGATCAATGCAAGAAATAGAGGAAAAACAACA 144877
 QY 1589 ----- 1588
 Db 144878 GAATGGGAAGACTAGGAGTCTCTTCAAGAAATACAGAGATACCAAGGAAACATTTTCATG 144937
 QY 1589 ----- 1588
 Db 144938 CAAGATGAGCTCGATNAAGGACAGAAATGGTATGGACCTAACAGAGACCAAGATATTA 144997
 QY 1589 ----- 1588
 Db 144998 AGAAGAGATGCCAAGATACACAGAGACCTGTACAAAAGATCTTCAAGACCCAGATA 145057
 QY 1589 ----- 1588
 Db 145058 ATCAGATGGTGTGATCACTCACTAGAGCCAGACATCCTGGAAATGTGAAGTCAAGTGGG 145117
 QY 1589 ----- 1588
 Db 145118 CTTAAGNAGCATCACTACGAAAGAGCTAGTGAGGTGATGGAATTCACAGTTTCAGCTAT 145177
 QY 1589 ----- 1588
 Db 145178 TCCAAATCCTGAAAGATGATGCTGTGAAAGTGTGCTGCACTCAATATGCCAAGAAATTTGGA 145237
 QY 1589 ----- 1588
 Db 145238 AAATCAGAGTGGCCACAGACTGGAAGAGTCAAGTCTTTTGTCCAAATCCCAAGAAAG 145297
 QY 1589 ----- 1588
 Db 145298 ACAATGACAAAGATGCTCAAGATACCGTACAAATTCGACTCATCTCACAGGCTAGTAAAG 145357
 QY 1589 ----- 1601
 Db 145358 TAATGCTCAAAATTTCTCAAGCCAGGCTTCAGCAATATGTGAACCGTGAACCTCCAGATG 145417
 QY 1602 TTCAAGCTGGTTTTAGAAAAGTCAAGAGAACAGAGACCAAAATGCGCAACATCTCTGTA 1661
 Db 145418 TTCAAGCTGGTTTTAGAAAAGTCAAGAGAACAGAGATCAAAATGCGCAACATCTGCTGGA 145477
 QY 1662 TCATGAAAAAGCAAGAGAGTTCCAGAAAAACATCTTTCTGCTTTATGACTATGCAA 1721
 Db 145478 TCATGAAAAAGCAAGAGAGTTCCAGAAAAACATCTTTCTGCTTTATGACTATGCAA 145537
 QY 1722 AAGCTTTGACTGTGGGGTCAAAATTAAGTGTGAAATTTCTGAAAGGATGGGAATAC 1781
 Db 145538 AAGCTTTGACTGTGGATCAAAATTAAGTGTGAAATTTCTGAAAGGATGGGAATAC 145597
 QY 1782 CAGACCACTGACTGACTCTTGAATAATTTGTATGAGTCAAGGACCAAGAGTTAGAA 1841
 Db 145598 CAGACCACTGACTGCTCTTTCAGAAATTTGTATGAGTCAAGGACCAAGAGTTAGAA 145657
 QY 1842 CTGAGATGAAACACAGACTGTTTCAGTAGGAAAGAGATGATCTCAAGGCTGTATAT 1901
 Db 145658 CTGAGATGAAACACAGACTGTTTCAGTAGGAAAGAGATGATCTCAAGGCTGTATAT 145717
 QY 1902 TGTCAACCGGCTGTTTAACTTCTATGACAG-ACATCATGAGAACGCTGGGCTGGAG 1960
 Db 145718 TGTCAACCGGCTGTTTAACTTCTATGACAG-ACATCATGAGAACGCTGGGCTGGAG 145777
 QY 1961 AAGCACAAGCTGGAATCAGATTGCGGAGAAATAGCAATACCTCAGATATGAGATG 2020
 Db 145778 AAGCACAAGCTGGAATCAGATTGCGGAGAAATAGCAATACCTCAGATATGAGATG 145837
 QY 2021 ATACCACCTTTATGGCAGAAAGTGAAGAGAACTAAGAGCCTCTTGTAGAGTGAAG 2080
 Db 145838 ACACCACTTTATGGCAGAAAGTGAAGAGAACTAAGAGCCTCTTGTAGAGTGAAG 145897
 QY 2081 AGGAGAGTGAAGAGTGGCTTAAAGCTCAACATTCAGAAAAACAGATCATGGATCTG 2140
 Db 145898 TAGAGAGTGAAGAGTGGCTTAAAGCTTAAATATTCAGAAAAACAGATCATGGATCG 145957

QY 2141 GTCCCATCACTTTCATGGGAAATAGATGGGAAACAGTGGAACAGTGTGACAGCTTTATTT 2200
 Db 145958 GTCCCATCACTTTCATGGGAAATAGATGGGAAACAGTGGAACAGTGTGACAGCTTTATTT 146017
 QY 2201 TTTGGGGGCTCCAAATCACTGACAGATGCTGACGACCATGAAATTTAAAAGACATT 2260
 Db 146018 TT--CTGGGCTCCAAATCACTACAGATGGTGAATGACGACCATGAAATTTAAAAGACATT 146075
 QY 2261 ACTCTCTGGAAGAAAGTTA--ACCAACTAGATAGCATATTTGAAAGCAGACGACATTAC 2318
 Db 146076 ACTCTCTGGAAGAAAGTTATGACCAACTAGATAGCATATTTCAAAGCAGACGACATTAC 146135
 QY 2319 CTTGCCAACAAAGCCCTCTAGTCAAGGCTATGCTTTTCCAGTGGTCAATGATGATG 2378
 Db 146136 TTTGCCAACAAAGATCCATCTAATCAAGGCTATGCTTTTCTCTGGTCAATGATGATG 146195
 QY 2379 TCAGAGTTGGACTGTGAAGAAAGCTGACGACCTGAAGAAATGATGCTTTTGAAGTGTGGT 2438
 Db 146196 TGAGAGTTGGACTGTGAAGAAAGGCTGACGCTGAAGAAATGATGCTTTTGAAGTGTGGT 146255
 QY 2439 TTGGAAGAGACTTTGAGAGTCCCTTGAGCTGCAAGGAGATCCCAACAGTCCATTTCTGAA 2498
 Db 146256 TTGCAAGAGACTTTGAGAGTCCCTTGAGCTGCAAGGAGATCCCAACAGTCCATTTCTGAA 146315
 QY 2499 GGAGATCAGCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAATCCAGTACTT 2558
 Db 146316 AGAGATCAGCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAATCCAGTACTT 146375
 QY 2559 TGGCCACCTGATCAGAGAGCTGACTCACTGGAAGAACCTGATGCTGGGAGGATGG 2618
 Db 146376 TGGCCACCTGATCAGAGAGTGTGACTCATTTGGAAGAAAGCTGATGCTGGGAGGATGG 146435
 QY 2619 GGGCAGGAGGAGAGGGGACGACAGAGATGAGATGGCTGGATGGATCACTGACTCGAT 2678
 Db 146436 GGGCAGGAGGAGAGGGGACGACAGA-GATGAGATGGCTGGATGGATCACTGACTCTAT 146494
 QY 2679 GGAGTGTGCTGGTGAACCTCTGAGTGGTGGTATGACAGGAGGCTGTCTGCGGC 2738
 Db 146495 GGAGTGTGCTGAGTGAACCTCGGAGTGGTATGACAGGAGGCTGTGCTGTGT 146554
 QY 2739 GATTGATGGGCTCAAGAGTGTGACACGACTGAGCAACTGAACTGAACTGAACTGTA 2797
 Db 146555 GATTGATGGGTTCCAAAGAGTCAAGAGTCAAGATGACTGAGTCACTGATCTGATGATATA 146613

RESULT 33 AC130787

LOCUS 173952 bp DNA linear HTG 27-SEP-2002
 DEFINITION Bos taurus clone RP42-206B24, WORKING DRAFT SEQUENCE, 3 ordered pieces.

AC130787
 AC130787.2 GI:23334708

VERSION HTG; HTGS PHASE2; HTGS_DRAFT.
 KEYWORDS Bos taurus (cow)

SOURCE Bos taurus

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

I (bases 1 to 173952)

Akhter, N., Antonellis, A., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blackley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Caraga, K., Coleman, B., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Loric, P., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schueler, M.G., Sison, C., Stantropop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Walker, M., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

NISC Comparative Sequencing Initiative
 Unpublished

TITLE
 JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

2 (bases 1 to 173952)
Green,E.D.
Direct Submission
Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 173952)
Green,E.D.
Direct Submission
Submitted (27-SEP-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Sep 27, 2002 this sequence version replaced gi:22218457.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: misc_zoo@nhgri.nih.gov
----- Project Information
Center project name: ddb
Center clone name: 206E24

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig, has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171764 bases at least Q40
Consensus quality: 172948 bases at least Q30
Consensus quality: 173503 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 173752; sum-of-contigs
Quality coverage: 5.74x in Q20 bases; agarose-fp
Quality coverage: 5.82x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

* 1 89789: contig of 89789 bp in length
* 89790 89889: gap of unknown length
* 89890 96759: contig of 8870 bp in length
* 96760 96859: gap of unknown length
* 96860 173952: contig of 77093 bp in length.

FEATURES
source

1. 173952
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-206E24"
/clone_lib="RP42"

misc_feature

1. 89789
/note="assembly_fragment
clone end.T7
vector_side:left"

misc_feature

89890..96759

/note="assembly_fragment"

misc_feature

96860..173952

/note="assembly_fragment"

missing SP6 clone_end from 3' end of insert"

Query Match	Best Local Similarity	Score	DB 1	Length
Matches 2536; Conservative 0; Mismatches 239; Indels 627; Gaps 16;	74.5%;	53.7%;	1919.1;	173952;
1	QTCAGGAAGGGCGGCAGTGGAGGAGGAGTACCTACCTCGTCCAAAGGTAAGGAGCAGTAGC 60	99257	GTCAGAGGGGCGAGCTGTGAGGAGA----	TACTGCTCGTCCAAAGTGTAGGAGCAGCAGC 99311
61	TGCGCTTTCTGGAGCAGCCCTAAAGAGATACCCCGCCCAAGGTAAAGAGAAACCCGAG 120	99312	TGCGCTTTCTGGAGCAGCCCTAAAGAGATACCCCGCCCAAGGTAAAGAGAAACCCGAG 99371	
121	TAAGATGTAGTGTGTGTGAGAGGCGCATCAGAGGCGGACATACCTAGTAAACCATACACGCA 180	99372	TAAGATGTAGTGTGTGTGAGA--GGCATCAGAGGCGGACACAC--AAACCATATACACA 99428	
181	GAAACTAGTCAATCTAATCATCATTAGGACACAGCCCTTGTCTAATCAATGAACCTAAG 240	99429	GAAACTAGTCAATCTAATCATCATTAGGACACAGCCCTTGTCTAATCTAGTGAACCTAAG 99488	
241	CGATGCC--CGTGGGGCAACCCAAAGATGGCGAGGTTCATGGTGGGAGATCTGACAGAAATG 298	99489	CGATGCC--CGTGGGGCAACCCAAAGATGGCGAGGTTCATGGTGGGAGATCTGACAGAAATG 99548	
299	TGCTCCACTGGAGAGGG--AATGCAAAACCACTTCAGTATTCTTGCCTTGAGAACCCCATG 357	99549	TGCTCCACTGGAGAGGGTAAAGGCAACCACTTCAGTATTCTTGCCTTGAGAACCCCATG 99608	
358	ACAGTATGAAAGCAAAATGATAGTATCTGAAGAGGAACTCCCGAGCTCAGTAGGT 417	99609	ACAGTATGAAAGCAAAATGATAGTATCTGAAGAGGAACTCCCGAGCTCAGTAGGT 99668	
418	GCCCCATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAG 477	99669	GCCCCATATGCTACTGGAGATCAGTGGAGAAATACTCCAGAGAGATGAAGAGGATGGAG 99728	
478	CCAAAGCAAAAGAAATACCCAGCTGTGGATGTGATGTGATGATTAAGAAAGTCCGATGC 537	99729	CCAAAGTGAACCAATACCCAGCTGTGGATGTGATGTGATGATTAAGAAAGTCCGATGC 99788	
538	TGTAAGAGCAATATTGCATAGGAACCTGGAATGTCAGGTCCATGAATCAAGGCAAAATG 597	99789	TGTAAGAGCAATATTGCATAGGAACCTGGAATGTCAGGTCCATGAATCAAGGCAAAATG 99848	
598	GAAAGTGGTCAAAACAAGA--GATGGCAAGAGTGAATGTCAACATTTAGGAATCAGCGAACT 656	99849	GAAATTAGTCAAAACAGGAGGATGGCAAGAGTGAACCTGCGACATTTCTAGGAATCAGCGAACT 99908	
657	AAATGGAGTGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTGCGGCA 716	99909	AAATGGAGTGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTACTGCGGCA 99968	
717	GGAAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAAGAGTCCGAAATGCACT 776	99969	GGAAATCCCTTAGAAGAAATGGAGTAGCCATCATGTCAACAAAAGAGTCCGAAATGCACT 100028	
777	ACTTGATGACTCTCAAAACGACAGAAATGATCTCTGTTTTCAGAGGCAAAACCAT 836	100029	ACTTGATGACTCTCAAAACGACAGAAATGATCTCTGTTTTCAGAGGCAAAACCAT 100088	
837	CAATATCAGTAATCCCAAGTCTATGCCCAACCAAGTAAATGCTGAAGAGCTGAAGTTGA 896	100089	CAATATCAGTAATCCCAAGTCTATGCCCAACCAAGTAAATGCTGAAGAGCTGAAGTTGA 100148	
897	ACGGTCTTATGAAGACCTTACAGACCTTTTGAAGCTAAACCCCAACCAAGAGTGTCTCT 956	100149	ATGGTCTTATGAAGACCTTACAGACCTTTTGAAGCTAAACCCCAACCAAGAGTGTCTCT 100208	
957	CATTATAGGGGACTGGAATGCAAAAGTAGGAGCAAGAAACACCTGGAGTAAACAGGCA 1016	100209	CATTATAGGGGACTGGAATGCAAAAGTAGGAGCAAGAAACACCTGGAGTAAACAGGCA 100267	

QY	1017	ATTGGCCCTTGGAAATACGAATGAAGCAGGGCAAGACTAATAGAGTTTGGCCAAGAAA	1076	QY	1591	-----	1590
Db	100268	ATGTGGCCCTTGGAAATACGAATGAAGCAGGGCAAGGCTA--ACAGTTTACCAACAGAA	100325	Db	101342	CCACACAATTGCACTCATCTCACACGCTGGTAAAGTAATGCTCAAAATCTCCCAAGCCAG	101401
QY	1077	TGCACTGTGTATAGCAAAACACCCCTCTTCCAAACAACAAGAGAGACTCTACACATGGAC	1136	QY	1591	-----	1590
Db	100326	TGCACGTGTATAGCAAAACACCCCTCTTCCAAACAACAAGAGAGACTCTACACATGGAC	100385	Db	101402	GCTTCTGCAATATGTGAACCAATGGACTTCCAGATGTTCAAGCTGTTTTCGAAAGGCGAG	101461
QY	1137	ATCACCAAGATGCTCAACACCGAATACAGATTGATTATATTCTTGGACGCCAAGATGGAG	1196	QY	1627	AGGAACCCAGAGACCAAAATGCAACATCTCTGTATCATGGAAGAACAA--GAGAGTTCC	1685
Db	100386	ATCACCAAGATGCTCAACACCGAATACAGATTGATTATATTCTTGGACGCCAAGATGGAG	100445	Db	101462	AGGAACCCAGAGATCAAAATGTCAACATCTCTGGATCATGGAAGAACAAAGAGAGATTCC	101521
QY	1197	AAGCTCTATACAGTCAGCAAAACCAAGCAGGAGCTTACTGTGGCTCAGATCATGAAC	1256	QY	1686	AGAAAAACATCTATTCTTGTCTTATTGACTATGCAAAAGCCCTTTGACTGTGGGGTCCACA	1745
Db	100446	AAGCTCTATACAGTCAGCAAAACCAAGCAGGAGCTTACTGTGGCTCAGATCATGAAC	100505	Db	101522	AGAAAAACATCTATTCTTGTCTTATTGACTATGCAAAAGCCCTTTGACTGTGGGGTCCACA	101581
QY	1257	CCTTATTTGCCAATTCAGACTTAATTTGAAGAAAGTAGGGAACCACTAGATCACTCAG	1316	QY	1746	ATAAACTGTGGAATTTCTGAAGGGATGGAAATACAGACCACTGACCTGACTCTTGA	1805
Db	100506	CCTTATTTGCCAATTCAGACTTAATTTGAAGAAAGTAGGGAACCACTAGATCACTCAG	100565	Db	101582	ATAAACTGTGGAATTTCTGAAGGGATGGAAATACAGACCACTGACCTGACTCTTGA	101641
QY	1317	GTAAAGACTTAATCCCAATCCCTTATGATTATACAGTGGAAAGTAGAATTAAGGG	1376	QY	1806	AAAAATTTGTATGCAAGTCAAGGAGCAACAGTTAGAACTTGGACATGGAACAACAGACTGGT	1865
Db	100566	GTAAAGACTTAATCCCAATCCCTTATGATTATACAGTGGAAAGTAGAATTAAGGG	100625	Db	101642	GAAACCTGTGTAGAGTCAAGGAGCAACAGTTAGAACTTGGACATGGAACAACAGACTGGT	101701
QY	1377	CCTAGATCTGTATAGACAGAGTA CCTAATGAATATGACAGAGGTTTATGACATTTGACA	1436	QY	1866	TCCAAAGTAGGAAAAAGGAGTATGTCAAGGCTGTATATTGTCTACCCGGCTTTGTTAACTTCT	1925
Db	100626	ACTAGATCTGTATAGACAGAGTA CCTAATGAATATGACAGAGGTTTATGACATTTGACA	100681	Db	101702	TCCAAATAGGAAAAAGGAGTATGTCAAGGCTGTATATTGTCTACCTGCTTATTTAAATAT	101761
QY	1437	GGACAGAGGATCGAGACCATCCCATGGAAGAAAGTAATGCAAAAGCAAAATGGCTGTC	1496	QY	1926	ATGCAGAG--ACATCATGAGAAACCGCTGGCTGGAGAGACACAGCTGGATCAAGATTG	1984
Db	100682	GGAGCAGAGATCAAGACCATCCCTAAGGAATAAATGCAAAAGCAAAATGGCTTAT	100741	Db	101762	ATGCAGATCATCATGAGAAACCGCTGGCTGGAGAGACACAGCTGGATCAAGATTG	101821
QY	1497	TGGGAGGCTTACAAATAGCTGTGAAAGAGAGAGTGAAGCAAGCAAGGAAAAAGGA	1556	QY	1985	CGGAGAGAAATAGCAATAACCTCAGATATGAGATGATACCAACCTTATGGCAAGAAAGTG	2044
Db	100742	TGGGAGGATTTACAAATAGCTGTGAAAGAGAGAGTGAAGCAAGCAAGGAAAAAGGA	100801	Db	101822	CGGAGAGAAATAGCAATAACCTCAGATATGAGATGATACCAACCTTATGGCAAGAAAGTG	2044
QY	1557	AAGTAAAGCATCTGAATCGAGTTCCAAAG	1590	QY	2045	AAGAGAACTTAAAGGCTCTTGTATGAAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAA	2104
Db	100802	AAGAAATTAAGCATCTGAATCGAGTTTCCAAAGAAATAGCAAGGAGAGATAAGAAAGCCCT	100861	Db	101882	AAGAGAACTTAAAGGCTCTTGTATGAAGTGAAGAGTGAAGAAAGTTGGCTTAA	101941
QY	1591	-----	1590	QY	2105	AGCTCAACATTCAGAAACCAAGATCATGGCATCTGGTCCCATCACTTTCATGGGAAATAG	2164
Db	100862	CCTCAGTGATCAATGCAAGAAATAGAGAAACAACTGATGGGAAACACTAGAGATCT	100921	Db	101942	AGCTTAACTTCAATAATGAAGTATGGCATCTGGTCCCATCACTTTCATGGGAAATAG	202001
QY	1591	-----	1590	QY	2165	ATGGGAAACAGTGGGAAACAGTGTGAGCTTTATTTTGGGGGGCTCCAAATCACTGC	2224
Db	100922	CTTCAAGAAAAATTAGAGATACGAAGGAAAAATTCATGCAAGATGGGCTCAGTAAAGGA	100981	Db	102002	ATGG-----GGAACAGTGTGAGCTTTA--TTTTGGGGGGCTCCAAATCACTGC	102050
QY	1591	-----	1590	QY	2225	AGATGGTCACTGCAAGCCATGAAATTTAAAGACACTTACTCTCTGGAAGAAAGTTA--AC	2282
Db	100982	CAGAAATGGTATGGACCTTAACAGACAGAGATATTAAGAGAGTGGCAAGATACAC	101041	Db	102051	AGATGGTCACTGCAAGCCATGAAATTTAAAGACACTTACTCTCTGGAAGAAAGTTA--AC	102110
QY	1591	-----	1590	QY	2283	CAACCTAGATAGCATATTGAAAGCAGAGACA--TTACCTTGCACCAAGAGGCTTCT	2339
Db	101042	AGAAGAACTGTACAAAAAAGCTTCAACCCAGATAATCAAGTGGTGTGATCACTCAT	101101	Db	102111	CACCTAGATAGCATATTGAAAGCAGAGAGACATATTTGCAACCAAGGTTCCATCT	102170
QY	1591	-----	1590	QY	2340	AGTCAAGCTATGTTTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG	2399
Db	101102	CTAGAGCCAGACATCCCTGAATGTGAATGATGATGATGATGATGATGATGATGATGATG	101161	Db	102171	AGTCAAGCTATGTTTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG	102230
QY	1591	-----	1590	QY	2400	AGCTGAGCAGTGAAGAAATGATGCTTTTGAACCTGTTGTTGTTGTTGTTGTTGTTGTTG	2459
Db	101162	AAGCTAGTGGAGTGTGAATTTCCAGTTTCACTATTTTCAATCTCTGGAAGATGATGCTT	101221	Db	102231	AGCTGAGTGGCGAAGAAATGATGCTTTTGAACCTGTTGTTGTTGTTGTTGTTGTTGTTG	102290
QY	1591	-----	1590	QY	2460	CCCTTGGACTGCAAGAGATCCACAGTGTGATGATGATGATGATGATGATGATGATGATG	2519
Db	101222	GAAAGTGTGCACTCAATATGCGAGCAAAATTTGGAAAACTCAGCAGTGGCCACAGACTG	101281	Db	102291	CCCTTGGACTGCAAGAGATCCACAGTGTGATGATGATGATGATGATGATGATGATGATG	102350
QY	1591	-----	1590	QY	2520	TTTGGAAAGGAATGATGCTTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCAAGAGC	2579
Db	101282	GAAAGTGTGCTTTTTCATTTCAATCCAAAGAAAGCAATGCTAAGAAATGCTCAACTA	101341	Db	102351	ATTGTTAGGACTGATGTTGAACTGAAATGCGAGTACTTTAGCCACCTGATGTGAAGAGC	102410
QY	101282	GAAAGTGTGCTTTTTCATTTCAATCCAAAGAAAGCAATGCTAAGAAATGCTCAACTA	101341	QY	2580	TGACTCACTGGAAGAACCTGATGCTGGAGGATTTGGGGGCGAGGAGGAGGAGGAGC	2639

```

Db 102411 TGACTCACAGAAAGACCCCTGATCTGGGAAGATTGGAGTAGGAGAGGGGATG 102470
QY 2640 ACAGAGATGAGATGGCTGATGGATCACTGACTCGATGGAGCTGAGTCTGGGTGAATC 2699
Db 102471 ACAGAGATGAGATGGTGGATGGATCACTCCGACTCAGTGGACACGGGTTCGGGTGACT 102530
QY 2700 CTGGAGTTGGTATGACACAGGAGGCGCTGCTCTCGGGCGATTTCATGGGGTTCACAAAGAG 2759
Db 102531 CTGGAGTTGATGATGACAGGAGGCGCTGCTCTGCTGATTCATGGGGTTCACAAAGAG 102590
QY 2760 TTGGACACGACTGAGCAACTGAATGAATGAATGAATGAATGAATGAATGAATGAATGA 2801
Db 102591 TCGGACAGGACTGAGCGACTGAAATGAATGAATGAATGAATGAATGAATGAATGA 102632

RESULT 34
AC092006
LOCUS Bos taurus clone RP42-456A17, WORKING DRAFT SEQUENCE.
DEFINITION AC092006
ACCESSION AC092006
VERSION 2 GI:21263249
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
REFERENCE 1 (bases 1 to 216698)
AUTHORS Akhtar,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Bresn,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B., Margulies,E.H., Masiello,C., Maekeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., McGuire,C., Pearson,R., Portnoy,N.E., Prasad,A., Schueler,M.G., Stantropop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Tsugeon,C., Vogt,J.B., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 216698)
Green,E.D.
Direct Submission
Submitted (13-JUN-2001) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 216698)
Green,E.D.
Direct Submission
Submitted (30-MAY-2002) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
On May 30, 2002 this sequence version replaced gi:14389328.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoehngri.nih.gov
----- Project Information
Center project name: cbj
Center clone name: 456A17

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

```

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 216492 bases at least Q40
Consensus quality: 216658 bases at least Q30
Consensus quality: 216689 bases at least Q20
Insert size: 187000; agarose-fp
Insert size: 216698; sum-of-contigs
Quality coverage: 14.77x in Q20 bases; agarose-fp
Quality coverage: 12.74x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 216698: contig of 216698 bp in length.
Location/Qualifiers
1. 216698
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="RP42-456A17"
/clone.lib="RP42"
misc_feature
1. 216698
/note="assembly_fragment
clone_end:SP6
vector_side:left
clone_end:T7
vector_side:right"
1. 42720
/note="clone overlaps with GenBank Accession Number
AC099746 clone RP42-244N18 (center project name crh)"
120825..216698
misc_feature
/note="clone overlaps with GenBank Accession Number
AC092007 clone RP42-73L22 (center project name cbk)"
Query Match 53.5%; Score 1909.81; DB 1; Length 216698;
Best Local Similarity 73.9%; Pred No. 0;
Matches 2473; Conservative 0; Mismatches 257; Indels 618; Gaps 10;
QY 282 AGAGATCTGACAGAAATGTGGTCCACTGGAGAGGGAAT-GCAACCACTTCAGTATTCTT 340
Db 17823 AGAGATCTGACAGAAATGTGGTCCACTGGAGAGGGAATGGCAACCACTTCAGTATTCTT 17882
QY 341 GCCTTGAGAACCCCATGACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAAC 400
Db 17883 GCCTTGAGAACCCCATGACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAAC 17942
QY 401 TCCCGAGGTTCAGTAGGTGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAA 460
Db 17943 TCCCGAGGTTCAGTAGGTGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAA 18002
QY 461 AGAATGAAGATGAGGCCAAAGCAAAAGATATCCAGCTGTGGATGTGACTGGTGATA 520
Db 18003 AGAATGAAGATGAGGCCAAAGCAAAAGATATCCAGCTGTGGATGTGACTGGTGATA 18062
QY 521 TAAGCAAGTCCGATGCTGTAAGAGCAATATTCATAGGAACCTTGGAAATGTCCAGTCCA 580
Db 18063 GAAGCAAGTCCGATGCTGTAAGAGCAATATTCATAGGAACCTTGGAAATGTCCAGTCCA 18122
QY 581 TGAATCAAGGCAAAATGGAAGTGGTCAACAAAGAGATGGCAAGATGAATGATCAACATTC 640
Db 18123 TGAATCAAGGCAAAATGGAAGTGGTCAACAAAGAGATGGCAAGATGAATGATCAACATTC 18182
QY 641 TAGGATCAGCAACTAAATGGAATGGGTAATTTAACTCAGATGACCAATTATA 700
Db 18183 TAGGATCAGCGGACTAAATGGAATGGGTAATTTAACTCAGATGACCGTTATA 18241

```

QY 701 TCTACTCTGGCGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAA 760
Db 18242 TCTACTCTGGCGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAA 18301
QY 761 GAGTCCGAAATGCAGTACTTGGATGCGAGTCTCAAAAAACGACAGAAATGATCTCTGTTGTT 820
Db 18302 GAGTCCGAAATGCAGTACTTGGATGCGAATCTCAAAAAATGACAGAAATGATCTCTGTTGTT 18361
QY 821 TCCAGGCAACCAATTCNAATACACAGTAATCCAGTCTATGCCCCAACAGTAATGCTG 860
Db 18362 TCCAGGCAACCAATTCNAATACACAGTAATCCAGTCTATGCCCCAACAGTAATGCTG 18421
QY 881 AAGAGGTGAAATGTAACGGTCTATGAAGACCTTACAGACCTTTTGAAGACTTAACACCCA 940
Db 18422 AAGAGGTGAAATGTAATGGTCTATGAAGACCTTACAGACCTTTTGAAGACTTAACACCCA 18481
QY 941 AAAAAGATGCTCTCTCAATATATAGGGGACTGGAATGCAAAAAGTAGGAAGCAAAAGAACAC 1000
Db 18482 AAAAAGATGCTCTCTCAATATATAGGGGACTGGAATGCAAAAAGTAGGAAGTCAAGAACAC 18541
QY 1001 CTGAGTAACAGGCAAAATTTGGCCCTTGGAAATACGGAATGAAGCAGGCAAGACTAATAG 1060
Db 18542 TTGAGTAACAGGCAAAATTTGG-CTTGGAAATACGGAATGAAGCAGGCAAGACTAATAG 18600
QY 1061 AGTTTGGCAAGAAATGCACTGCTCATAGCAAAACACCTCTTCCAAACACACAGAGAA 1120
Db 18601 AGTTTGGCAAGAAATGCACTGCTCATAGCAAAACACCTCTTCCAAACACACAGAGAA 18660
QY 1121 GACTCTACAGATGACATCACCAGATGGTCAACACGAAATCAGATTGATTAATCTTT 1180
Db 18661 GACTCTACAGATGACATCACCAGATGGTCAACACGAAATCAGATTGATTAATCTTT 18720
QY 1181 GCAGCCAAAGATGAGAGAGTCTATACAGTCAGCAAAACACAGACCAAGAGAGTACTGTG 1240
Db 18721 GCAGCCAAAGATGAGAGAGTCTATACAGTCAGCAAAACACAGACCAAGAGAGTACTGTG 18780
QY 1241 GCTCAGATCATGAATCTCTTATGCCCAAATTCAGACTTAATTAAGTGAAGAGAGAGAA 1300
Db 18781 GCTCAGATCATGAATCTCTTATGCCCAAATTCAGACTTAATTAAGTGAAGAGAGAGAA 18840
QY 1301 CCAGTAGATCACTCAGGTAAAGCTTAATCCATCCCTTATGATTAATACAGTGCAGAGTGA 1360
Db 18841 CCAGTAGATCACTCAGGTAAAGCTTAATCCATCCCTTATGATTAATACAGTGCAGAGTGA 18900
QY 1361 GAAATAGATTAAGGCTAGATCTGATGACAGAGTACTTAATGAACCTATGACAGAGAGG 1420
Db 18901 GAAATAGATTAAGGCTAGATCTGATGACAGAGTACTTAATGAACCTATGATGAGAGG 18960
QY 1421 TTCATGACATTTACAGGACAGGAGTCCAGACCTCCCATGGAAAGAAATGCAAAA 1480
Db 18961 TTCATGACATTTACAGGACAGGAGTCCAGACCTCCCATGGAAAGAAATGCAAAA 19020
QY 1481 AAGCAAAATGGCTGCTGGGAGGCTTACAAATAGCTGTGAAAGAGAGAGTGAAGAA 1540
Db 19021 AAGCAAAATGGCTGCTGGGAGGCTTACAAATAGCTGTGAAAGAGAGAGTGAAGAA 19080
QY 1541 GCAGGAAAGAAAGGAAAGTAAAGCATCTGATGCAAGTTCCTGAGTTCCTGAGTTCCTGAG 1588
Db 19081 GCAGGAAAGAAAGGAAAGTAAAGCATCTGATGCAAGTTCCTGAGTTCCTGAGTTCCTGAG 19140
QY 1589 ----- 1588
Db 19141 GCGATAAGAAAGCTTCTTCAGCGATCAATGCAAGAAATAGAGGAAACACAGAGATGG 19200
QY 1589 ----- 1588
Db 19201 GAAAGACTAGAGATCTCTCAAGAAATCAGAGATACCAAGGAAACATTTTCATGCAAGA 19260
QY 1589 ----- 1588
Db 19261 TGGGCTCGATAAGGACAGAAATGGTATGGACCTTAACAGNAGCAGAGATATTAAGAAGA 19320

QY 1589 ----- 1588
Db 19321 GGTGGCAAGAAATACACAGAGAACTGTACAAAAAGATCTTCATGACCCAGATATCAAG 19380
QY 1589 ----- 1588
Db 19381 ATGGTGTGATCACTGACCTAGAGCCAGACATCCTGGAATGTGAAGTCAAGTGGGCTTAG 19440
QY 1589 ----- 1588
Db 19441 AAAGCATCACTACGAAACAAAGCTAGTGAGGTGATGGAATTCAGCTGAGCTCTTTTCAA 19500
QY 1589 ----- 1588
Db 19501 TCTTAAAGATGACGCTGTGAAGTGTGCACTCAATATATGACCAAAATTTGGAAACATC 19560
QY 1589 ----- 1588
Db 19561 AGCAGTGGCCACAGGACTGGAAAGGTGAGTTTCATTCATTCCTCAATCCCAAGAAAGCAATG 19620
QY 1589 ----- 1588
Db 19621 CCAGAAATGCTCAAACTACCGCACAGTTGCATCTCATGCTAGTAAAGTAGTGC 19680
QY 1589 ----- 1607
Db 19681 TCAAAATTTCTCAAGCCAGGCTTCAGCAATATGTGAACCGTGAACCTTCTGATGTTCAA 19740
QY 1608 CTGGTTTGTAGAAAGTTCAGAGAAACCCAGACCAAAATTTGCCAATCTCTGATCATG 1667
Db 19741 CTGGTTTGTAGAAAGTTCAGAGAAACCCAGAGATCGAATTTGCCAAGCTCTGATGATCATG 19800
QY 1668 AAAAGCAAGAGAGTTCAGAAAAACATCTATTTCTGCTTTATTTGACTATGCAAAAGCCT 1727
Db 19801 AAAAGCAAGAGAGTTCAGAAAAACATCTATTTCTGCTTTATTTGACTATGCAAAAGCCT 19860
QY 1728 TTGACTGTGGGGTTCAGAAATTAACCTGTGGAATTTCTGAAAGGATGGATACAGAC 1787
Db 19861 TTGATGTGTAGATCACAATCACTGTGGAATTTCTGAAAGAGATGGGATACAGAC 19920
QY 1788 ACCTGACCTGCTTCTGAAAAATTTGTATGAGGTGAGAAACCAAGTTAGAACTGGAC 1847
Db 19921 ACCTGATCTGCTCTTGAGAAATTTGTATGAGGTGAGAAACCAAGTTAGAACTGGAC 19980
QY 1848 ATGGAACAAACAGACTGGTTCCAGTAGGAAGAGATATGTCAAGGCTGTATATGTCAC 1907
Db 19981 ATGGACAAACAGACTGGTTCCAAATAGGAAAGAGTACGTCAAGGCTGTATATATCAC 20040
QY 1908 CCGGCTTCTTTAACTTCTATGCAGAG-ACATCATGAGAAACGCTGGGCTGGAAGAACAC 1966
Db 20041 CCGTGTATTTAACTTATGACAGATACATCATGAGAAATGCTGGACTGGAAGAACAC 20100
QY 1967 AAGTGGAAATCAAGATTGCGGGAGAAATAGCAATAACCTCAGATATGCAAGATGATACCA 2036
Db 20101 AAGTGGAAATCAAGATTGCGGGAGAAATAGCAATAACCTCAGATATGCAAGATGATACCA 20160
QY 2027 CCGTTATGGCAGAAAGTGAAGAGGACTTAAAGGCTCTTTGATCAAGGTGAAAGAGAGAG 2086
Db 20161 CCGTTATGACAGAAAGTGAAGAGGACTTAAAGGCTCTTTGATCAAGGTGAAAGAGAGAG 20220
QY 2087 GT-GAAAAAGTTGGCTTAAAGCTCAACATTCAGAAAAACGAGATCATGGCATCTGGTCCC 2145
Db 20221 GTGGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAACGAGATCATGGCATCTGGTCCC 20280
QY 2146 ATCACTTCATGGGAAATAGATGGGAAACAGTGGAAACAGTGCAGACTTTATTTTGG 2205
Db 20281 ATCACTTCATGGGAAATAGATGGGAAACAGTGGAAACAGTGCAGACTTTATTTTGG 20338
QY 2206 GGGGCTCCAAATCACTGCAGATGGTGCAGCCATGAAATTTAAAGACACTTTACTCC 2265
Db 20339 GGGGCTCCAAATCACTGCAGATGGTGCAGCCATGAAATTTAAAGACACTTTACTCC 20398
QY 2266 TTGGAAGAAAAAGTTA--ACCACTTAGATGATATTTGAAAAAGCAGAGACATTTACCTTGC 2323

[illegible]

RESULT 35
AC099746
LOCUS AC099746
DEFINITION Bos taurus clone R242-244N18, WORKING DRAFT SQUENCE, 4 ordered pieces.
AC099746 171875 bp DNA linear HTG 30-MAY-2002
Bos taurus
AC099746.2 GI:21263247
HTG; HTGS_PHASE2; HTGS_DRAFT.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.
1 (bases 1 to 171875)
REFERENCE
Akhtar, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Leric, P., Lee-Hin, S.-Q., Legasp, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pacuigian, C., Pearson, R.

Portnoy, M.E., Prasad, A., Schueler, M.G., Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsurgenev, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 171875)
Direct Submission
Green, E.D.
Submitted (20-NOV-2001) NIH Intramural Sequencing Center, 8717
Grovmont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 171875)
Direct Submission
Green, E.D.
Submitted (30-MAY-2002) NIH Intramural Sequencing Center, 8717
Grovmont Circle, Gaithersburg, MD 20877, USA
On May 30, 2002 this sequence version replaced gi:17017550.

----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc_zco@hqi.nih.gov
 ----- Project Information
 Center project name: crh
 Center clone name: 244N18

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171182 bases at least Q40
Consensus quality: 171457 bases at least Q30
Consensus quality: 171549 bases at least Q20
Insert size: 146000; agarose-fp
Insert size: 171575; sum-of-configs
Quality coverage: 12.65x in Q20 bases; agarose-fp
Quality coverage: 10.76x in Q20 bases; sum-of-configs

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 123482: contig of 123482 bp in length
* 123483 123582: gap of unknown length
* 123583 137342: contig of 13760 bp in length
* 137343 137442: gap of unknown length
* 137443 142844: contig of 5402 bp in length
* 142845 142944: gap of unknown length
* 142945 171875: contig of 28931 bp in length.

```

```

FEATURES
* 142945 1/1875: config of
Location/Qualifiers
source 1..171875
        /organism="Bos taurus"
        /mol_type="genomic DNA"
        /db_xref="taxon:9913"
        /clone="RP42-244N18"
        /clone_lib="RP42"
misc feature 1..123382

```

/note="assembly_fragment clone_end:r7 vector_side:left 1..22552 /note="clone overlaps with GenBank Accession Number AC099407 clone RP42-270J1 (center project name crg)" 123583..137342 /note="assembly_fragment" 129328..171875 /note="clone overlaps with GenBank Accession Number AC092006 clone RP42-456A17 (center project name cbj)" 137443..142844 /note="assembly_fragment" 142945..171875 /note="assembly_fragment clone_end:SP6 vector_side:right"		Query Match Best Local Similarity 53.4%; Score 1907.01; DB 1; Length 171875; Matches 2499; Conservative 0; Mismatches 300; Indels 624; Gaps 10;	
misc_feature	282	AGAGATCTGACAGAAATGGTCTGACAGAGAGGGAAT-GCAAAACCACTTCAGTATTCCT	340
misc_feature	146977	AGAGATCTGACAGAAATGGTCTGACAGAGAGGGAATGGCAACCACTTCAGTATTCCT	147036
misc_feature	341	GCCTTGAGAACCCATGAACAGTATGAAAGGCCAAATGATAGGATCTGAAAGAGGAAC	400
misc_feature	147037	GCCTTGAGAACCCATGATAGTACGAAAGGCCAAATGATAGGATCTGAAAGAGGAAC	147096
misc_feature	401	TCCCAGGTCACTAGTGGTCCCATATGCTACTGGAGATCTGAGTGGAGAAATAACTCCAGAA	460
misc_feature	147097	TCCCAGGTCACTAGTGGTCCCATATGCTACTGGAGATCTGAGTGGAGAAATAACTCCAGAA	147156
misc_feature	461	AGATGAGAGATGGCCCAAGCAAAAGATACCCAGCTGTGGATGTGACTGGTGATA	520
misc_feature	147157	AGATGAGAGATGGCCCAAGCAAAAGATACCCAGCTGTGGATGTGACTGGTGATA	147216
misc_feature	521	TAAGCAAGGTCCGATGCTGTTAAAGAGCAATATTCATAGGAACCTGGAATGTGAGTCCA	580
misc_feature	147217	GAAGCAAGGTCCGATGCTGTTAAAGAGCAATATTCATAGGAACCTGGAATGTGAGTCCA	147276
misc_feature	591	TGAATCAAGGCAATTTGGAATGTTGTAACAGAGATGGCAGAGTGAATGTCAACATTC	640
misc_feature	147277	TGAATCAAGGCAATTTGGAATGTTGTAACAGAGATGGCAGAGTGAATGTCAACATTC	147336
misc_feature	641	TAGGAATCAGCGAATCAAAATGGACTGGAATGGGTGAAATTTAACTCAGATGACCAATTATA	700
misc_feature	147337	TAGGAATCAGCGAATCAAAATGGACTGGAATGGGTGAAATTTAACTCAGATGACCGTTATA	147395
misc_feature	701	TCTACTACTGGGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAA	760
misc_feature	147396	TCTACTACTGGGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAA	147455
misc_feature	761	GAGTCCGAATTCGACTACTTCGATGCGCTCTCAAAAGCAGAGATGATCTCTGTTTGT	820
misc_feature	147456	GAGTCCGAATTCGACTACTTCGATGCGCTCTCAAAAGCAGAGATGATCTCTGTTTGT	147515
misc_feature	821	TCCAGGCAAAACCAATTCATATCAAGTAATCCAAAGTCTATGCCCCAACCAAGTAATGCTG	880
misc_feature	147516	TCCAGGCAAAACCAATTCATATCAAGTAATCCAAAGTCTATGCCCCAACCAAGTAATGCTG	147575
misc_feature	881	AAGAAGCTGAAGTTGAACGGTCTTATGAGACCTTACAGACCTTTTAGACTACACCCA	940
misc_feature	147576	AAGAAGCTGAAGTTGAACGGTCTTATGAGACCTTACAGACCTTTTAGACTACACCCA	147635
misc_feature	941	AAAAAGATGCTCTTCTCATTTAGGGACTGGAATGCAAAAGTAGGAAGCAAAAGAACAC	1000
misc_feature	147636	AAAAAGATGCTCTTCTCATTTAGGGACTGGAATGCAAAAGTAGGAAGTCAAGAAACAC	147695
misc_feature	1001	CTGGAGTAACAGGCAAAATTTGGCCCTTGGAAATACGGAATGAGCAGGCGCAAGACTATAG	1060
misc_feature	147696	TTGGAGTAACAGGCAAAATTTGGCCCTTGGAAATACGGAATGAGCAGGCGCAAGACTATAG	147755

QY	1061	AGTTTTGCCAAGAAATGCACCTGGTCAATAGCAAAACACCCCTCTTCCAAACACACAGAGAA	1120
Db	147756	AGTTTTGCCAAGAAATGCACCTGGTCAATAGCAAAACACCCCTCTTCCAAACACACAGAGAA	147815
QY	1121	GACTCTACACATGGACATCACAGATGGTCAACACCGAAATCAGATTGATTATTTCTTT	1180
Db	147816	GACTCTACACATGGACATCACAGATGGTCAACACCGAAATCAGATTGATTATTTCTTT	147875
QY	1181	GCAGCAAGATGGAGAGCTCTATACAGTCAAGCAAAACCAAGACCGAGCTTACTGTG	1240
Db	147876	GCAGCAAGATGGAGAGCTCTATACAGTCAAGCAAAACCAAGACCGAGCTTACTGTG	147935
QY	1241	GCTCAGATCATGAACCTCTTATTTGCCAAATTCAGACTTTAAATTTGAAGAAATGAGGAAAA	1300
Db	147936	GCTCAGATCATGAACCTCTTATTTGCCAAATTCAGACTTTAAATTTGAAGAAATGAGGAAAA	147995
QY	1301	CCACTAGATCACTCAGGTAAGACCTAAATCCATCCCTTATGATTATATACAGTGGAGTGA	1360
Db	147996	CCACTAGATCACTCAGGTAAGACCTAAATCCATCCCTTATGATTATATACAGTGGAGTGA	148055
QY	1361	GAATAGATTTAAGGCCCTAGATCTGATAGACAGAGTACCTAATGAACATATGGACAGG	1420
Db	148056	GAATAGATTTAAGGCCCTAGATCTGATAGACAGAGTACCTAATGAACATATGGACAGG	148115
QY	1421	TTTCATGACATTTGACAGAGACAGGATCGAGACCATCCCATGGAAGAAATGCAAAA	1480
Db	148116	TTTCATGACATTTGACAGAGACAGGATCGAGACCATCCCATGGAAGAAATGCAAAA	148175
QY	1481	AAGCAAAATGGTGTCTCGGGAGGCCCTTACAAATAGCTGTGAAAGAAAGAGAGTGAATA	1540
Db	148176	AAGCAAAATGGTGTCTCGGGAGGCCCTTACAAATAGCTGTGAAAGAAAGAGAGTGAATA	148235
QY	1541	GCAAGGAAAAAGGAAAGATAAAAGCATCTGAATGACAGTCTCAAA	1588
Db	148236	GCAAGGAAAAAGGAAAGATAAAAGCATCTGAATGACAGTCTCAAAAGAAATAGCAAGAA	148295
QY	1589	-----	1588
Db	148296	GCAATAAGAAAGCCCTTCTTCAGCGATCAATGCAAGAAATAGAGAAACACAGATGG	148355
QY	1589	-----	1588
Db	148356	GAAAGACTAGAGATCTCTTCAAGAAATACAGAGATACCAAGGAACATTTTCATGCAAGA	148415
QY	1589	-----	1588
Db	148416	TGGGCTCGATAAAGCAGAGAAATGGTATGGACCTAACAGAGCAGAGATATTAGAAAGA	148475
QY	1589	-----	1588
Db	148476	GGTGGCAAGATACACAGAGAACTGTACAGAAAGATCTTCAGCCAGATATACAG	148535
QY	1589	-----	1588
Db	148536	ATGGTGTGATCACTGACCTAGAGCCAGACATCCTGGAATGTGAAGTCAAGTGGGCTTAG	148595
QY	1589	-----	1588
Db	148596	AAAGCATCACTAGCAACAAAGCTAGTGGAGGTGATGGAAATTCAGCTTCTTCAAA	148655
QY	1589	-----	1588
Db	148656	TCTTAAAGATGATGCTGTGAAAGTGTGCTCACTCAATATGCCAGCAAAATTTGAAAACTC	148715
QY	1589	-----	1588
Db	148716	AGCAGTGGCCACAGGACTGGAAAAAGTCAAGTTTTCATTCCAATCCCAAGAAACCAATG	148775
QY	1589	-----	1588
Db	148776	CCAAAGAAATGCTCAAACTACCACACAGATTGCACATCTCATCTGCTAGTAAAGTAGTGC	148835

```
QY 1589 -----GAACTCCAGTGTTCAG 1607
Db 148936 TC AAAATTC CCAAGCAGCGCTTCAGCAATATGTGAACCGTGAACCTTCCTGATGTTCAAG 148995
QY 1608 CTGGTTTTAGAAAAGTCAGAGGAACACAGAGACCAAAATTTGGCAACATCTCTGTATCATGG 1667
Db 148996 CTGGTTTTAGAAAAGTCAGAGGAACACAGAGATCGAATTTGCCAAGCTCTGTATCATTTG 148955
QY 1668 AAAAGCAAGAGATTCAGAAAACATCTATTCTGCTTTATGACTATGCAAAAGCCT 1727
Db 148956 AAAAGCAAGAGATTCAGAAAACATCTATTCTGCTTTATGACTATGCAAAAGCCT 149015
QY 1728 TTGACTGTGGGGTCACAAATAAATCTGGGAAATCTTGAAGGAGTGGGAATACCAAGACC 1787
Db 149016 TTGATGTGTAGATCAACAATCTGGAATTTCTGAAGAGATGGGATACCAAGACC 149075
QY 1788 ACTGACCTGACTCTTGA AAAATTTGTATGAGGTCAAGGACAAAGATTTAGAACTGGAC 1847
Db 149076 ACTGATCTGCCTCTTGAGAAAATTTGTATGAGGTCAAGGACAAAGATTTAGAACTGGAC 149135
QY 1848 ATGGAACAACAGACTGTTCCAAAGTAGAAAAGGATATGTCGAAGCTGTATATCTGCAC 1907
Db 149136 ATGGACAACAGACTGTTCCAAATAGAAAAGGATATGTCGAAGCTGTATATCTGCAC 149195
QY 1908 CCGCTTTGTTTAACTTCTATGCAGAG--ACATCATGAGAAAACGCTGGGCTGGAAGAACAC 1966
Db 149196 CCGCTTTTAACTTATATGCAGAGTACATCATGAGAAATGCTGAGCTGGAAGAACAC 149255
QY 1967 AAGCTGGAATCAAGATTCGCGGAGAAATAGCAATAACCTCAGATATGAGATGATACCA 2026
Db 149256 AAGCTGGAATCAAGATTCGCGGAGAAATATCAATAACCTCAGATATGAGATGATACCA 149315
QY 2027 CCCTTATGGCAGAAAGTGAAGAGGAACTTAAAGACCTCTTTGATGAAGGTGAAGAGAGAGA 2086
Db 149316 CCCTTATGAGAGATGAAGAGAACTCAAAGCCCTTTGATGAAGGTGAAGAGAGAGA 149375
QY 2087 GT-GAAAAGTTGCTTAAAGCTCAACATTCAGAAAACGAAGATCATGGATCTGTGCCC 2145
Db 149376 GTGGAAAAGTTGCTTAAAGCTCAACATTCAGAAAACGAAGATCATGGATCTGTGCCC 149435
QY 2146 ATCACTTTCATGGGAAATAGATGGGAAACAGTGAACACGTGTCAGACTTTATTTTGG 2205
Db 149436 ATCACTTTCATGGGAAATAGATGGGAAACAGTGAACACGTGTCAGACTTTA--TTTGG 149493
QY 2206 GGGCTCCAAAATCACTGCGAGATGGTGAATGCGACCATGAAATTAAGACACTTACTCC 2265
Db 149494 GGGCTCCAAAATCACTGCGAGATGGTGAATGCGACCATGAAATTAAGAGATGCTTACTCC 149553
QY 2266 TTGGAAGAAAAGTTA--ACCAACTAGATGATGATTTGAAGAGAGACATTAACCTTGC 2323
Db 149554 TTGGAAGAAAATTTATGACCAATCTAGATGATGATTTCAAAGAGAGACATTAACCTTGC 149613
QY 2324 CAACAAAGCCCCATCTAGTCAAGCTATGGTTTTTCCAGTGGTCAATGATGGATGTGAGA 2383
Db 149614 CAACAAAGGTCATCTAGTCAAGCTATGGTTTTTCCAGTGGTCAATGATGGATGTGAGA 149673
QY 2384 GTTGGACTGTGAAGAAAGCTGAGACACTGAAGAAATGATGCTTTTGAATGCTGTGTGGA 2443
Db 149674 GTTGGACTGTGAAGAAAGCTGAGTGCCTGAAGAAATGATGCTTTTGAATGCTGTGTGGA 149733
QY 2444 GAAGACTCTTGAGAGTCCCTTGGAGTCAAGGAGATCCAAACCACTCCATTCGAAGAGAGA 2503
Db 149734 GAAGACTCTTGAGAGTCCCTTGGAGTCAAGGAGATCCAAACCACTCCATTCGAAGAGAGA 149793
QY 2504 TCAGCCCTGGGATTTCTTGGAAAGGAATGATGCTTAAAGCTGAACTCCAGTACTTTGGCC 2563
Db 149794 TCAGCCCTGGGATTTCTTGGAAAGGAATGATGCTTAAAGCTGAACTCCAGTACTTTGGCC 149853
QY 2564 ACCTGATCAGNAGAGTGAATCTGCAAGAAAGACCTGATGCTGGGAGGATTTGGGGCA 2623
Db 149854 ACCTCATGCGAAGAGTTGACTCATTTGGGAAAGACTCTGATGCTGGGAGGATTTGGGGCA 149913
QY 2624 GGAGGAGAAAGGGACGACAGAGGATGAGATGGCTGGATGCGATCACTGACTCGATGAGC 2683
```

```
Db 149914 AGAGGAGAAAGCGGACGACAGAGGATGAGATGGCTGGATGGCATCACTGACTCGAT-GAGC 149972
QY 2684 TGAGTCTGGGTGAACCTCCTGGAGTTGGTATGAGCAGGGAGGCGCTGTCTCTGGGGGATTC 2743
Db 149973 TGAGTCTGAGTGAACCTCCTGGAGTTGGTATGAGCAGGGAGGCGCTGTCTGGGATTC 150032
QY 2744 ATGGGGTCCACAAGAGTTGGGACGAGCTGAGCACTGAGCACTGAACTGAACTGACTGAA 2803
Db 150033 ACGGGGTCAAAAGAGTCGGACACGAGCTGAGCTGATCTGATCTGATGTG----- 150084
QY 2804 CCTTAGTAGTTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2863
Db 150085 -----AAAGTTTATTTACAACGTTGTTAGTTTTCAGGTTGACAAAGATGATTCAGT 150137
QY 2864 ATGTTGGTTAAGATAAAGATTTTCAAATGATTTTATTTATTTTATTTTCTTCTACTT 2923
Db 150138 CATACATATACATACATCAITCTTTCTCAAATTTCTTTCTCATATAGGTTTTCACGAA 150197
QY 2924 ATTTAAATTTGGGATTTTAACTTAATTTCTTCAATGACTTG-----TATTTCTAATAATTT 2976
Db 150198 TATTGAATATAGGACCTGTGATATTCAGTAGGTTCTTGTTCATTAATCTGCTTACATAT 150257
QY 2977 ACTTATCTTATTTACTTTAATTAATTAATTAATTTTATTTATTTTCTTAATAAATCCAGT 3036
Db 150258 AGTAGTGTGTATAATCTAGGAAATCTTTTATTTGTTGGCTGTATAGTAAAAATTTAT 150317
QY 3037 CCTTCTTTTTTAAAGACCTTTAAATTAATTAATTTCTTTAGTGTTTTACCAGTTCT 3096
Db 150318 CATAGACTGGGTAAGAAATGAATCTCTTCTAGCTTAAGATTTGCTATAAATTTCT 150377
QY 3097 TTC 3099
Db 150378 TCC 150380

RESULT 36
AC134960
LOCUS
DEFINITION AC134960 183048 bp DNA linear HTG 03-DEC-2002
Bos taurus clone RP42-246J24, WORKING DRAFT SEQUENCE, 4 ordered
pieces.
ACCESSION AC134960
VERSION AC134960.2 GI:26005568
KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 183048)
REFERENCE
AUTHORS Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Marulles,E.H., Masello,C., Maskeri,B., McDowell,J.,
Paguirigan,C., Pearson,K., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantropop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.B.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 183048)
Green,E.D.
Direct Submission
Submitted (03-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 183048)
Green,E.D.
Direct Submission
Submitted (03-DEC-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Dec 3, 2002 this sequence version replaced gi:23477845.
```

----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@ngri.nih.gov
 ----- Project Information
 Center project name: djx
 Center clone name: 246J24

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 182395 bases at least Q40
 Consensus quality: 182646 bases at least Q30
 Consensus quality: 182726 bases at least Q20
 Insert size: 140000; agarose-fp
 Insert size: 182748; sum-of-contigs
 Quality coverage: 10.60x in Q20 bases; agarose-fp
 Quality coverage: 8.12x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 1 49758: contig of 49758 bp in length
 49759 49858: gap of unknown length
 49859 120040: contig of 70182 bp in length
 120041 120140: gap of unknown length
 120141 169454: contig of 49314 bp in length
 169455 169554: gap of unknown length
 169555 183048: contig of 13494 bp in length.

FEATURES

Source

1..183048
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-246J24"
 /clone_lib="RP42"

misc_feature

1..49758
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"

misc_feature

49859..120040
 /note="assembly_fragment"
 120141..169454

misc_feature

/note="assembly_fragment"
 169555..183048
 /note="assembly_fragment
 clone_end:T7
 vector_side:right"

Query Match

Best Local Similarity 52.9%; Score 1891.1; DB 1; Length 183048;
 Matches 2361; Conservative 0; Mismatches 164; Indels 603; Gaps 7;

QY	272	GTCTGTTGGAGAGATCTGACAGAAATGTGGTCCACTGGAGAGGGAAT-GCAAACCACTT	330
DB	41609	GTTTAGATAGACAGTTTTCACAGAAATGTGGTCCACTGGAGAGGGAATGGCAAACCACTT	41668
QY	331	CAGTATTTCTTGGCTTGAGAAACCCCATGAAAGATGAAAGGCAAAATGATAGATCTG	390
DB	41669	AAGTATTTCTTGGCTTGAGAAACCCCATGAAAGATGAAAGGCAAAATGATAGATCTG	41728
QY	391	AAAGAGGAATCCCCAGGTCAAGTGTGCTTACTGAGATCACTGGAGAAAT	450
DB	41729	AAAGAGGAATCCCCAGGTCAAGTGTGCTTACTGAGATCACTGGAGAAAT	41788
QY	451	AACTCCAGAAAGATGAAGAGATGAGGCAAGCAAAAGAAATACCCAGCTGTGGATGTG	510
DB	41789	AACTCCAGAAAGATGAAGAGATGAGGCAAGCAAAAGAAATACCCAGCTGTGGATGTG	41848
QY	511	ACTGTGTATATAAGCAAGGTCCGATGCTGTAAAGAGCAATATTGATAGGAACCTGGAAT	570
DB	41849	ACTGTGTATATAAGCAAGGTCCGATGCTGTAAAGAGCAATATTGATAGGAACCTGGAAT	41908
QY	571	GTCAAGTCCATGAATCAAGGCAAAATTTGAACTGTGTCAAACAGAGATGGCAAGAGTGAAT	630
DB	41909	GTCAAGTCCATGAATCAAGGCAAAATTTGAACTGTGTCAAACAGAGATGGCAAGAGTGAAT	41968
QY	631	GTCAACATTTAGGATCAGCAAGTAAATGGATGGATGGATGGATGGATGGATGGATGGAT	690
DB	41969	GTCAACATTTAGGATCAGCAAGTAAATGGATGGATGGATGGATGGATGGATGGATGGAT	42028
QY	691	GACCAATTATCTACTCTGCGGCGAGGAATCCCTCAGAAGAAATGGAGTAGCCATCATG	750
DB	42029	GACCAATTATCTACTCTGCGGCGAGGAATCCCTCAGAAGAAATGGAGTAGCCATCATG	42088
QY	751	GTCAACAAAGAGTCCGAAATTCAGTACTTGGATGGATGGATGGATGGATGGATGGATGGAT	810
DB	42089	GTCAACAAAGAGTCCGAAATTCAGTACTTGGATGGATGGATGGATGGATGGATGGATGGAT	42148
QY	811	TCGTGTTGTTTCCAAAGGCAAAACCAATTCAGTAAATCCAAAGTCTATGCCCCCAACC	870
DB	42149	TCGTGTTGTTTCCAAAGGCAAAACCAATTCAGTAAATCCAAAGTCTATGCCCCCAACC	42208
QY	871	-AGTAACTGCTGAAGAGTGAAGTGAACGGTCCCTATGAAGACCTACAGACCTTTTAGA	929
DB	42209	-AGTAACTGCTGAAGAGTGAAGTGAACGGTCCCTATGAAGACCTACAGACCTTTTAGA	42268
QY	930	ACTAACACCCCAAAAGAGTGTCTTCTCATTTATAGGGGACCTGGAATGCAGAAAGTAGAAG	989
DB	42269	ACTAACACCCCAAAAGAGTGTCTTCTCATTTATAGGGGACCTGGAATGCAGAAAGTAGAAG	42328
QY	990	CAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCTTTGGAAATACGGAATGAAGCAGGGCA	1049
DB	42329	CAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCTTTGGAAATACGGAATGAAGCAGGGCA	42388
QY	1050	AAGACTAATAGAGTTTCCCAAGAAATGCACTGGTCTATAGCAACACCTCTTCCCAACA	1109
DB	42389	AAGACTAATAGAGTTTCCCAAGAAATGCACTGGTCTATAGCAACACCTCTTCCCAACA	42448
QY	1110	ACACAAGAGAGAGCTCTACACATGCACATCACAGATGGTCAACACCCGAAATCAGATTGA	1169
DB	42449	ACACAAGAGAGAGCTCTACACATGCACATCACAGATGGTCAACACCCGAAATCAGATTGA	42508
QY	1170	TTATATTCTTTGGACCAAGATGGAGAGCTCTATACAGTCAGCAAAACCAAGACCAAGG	1229
DB	42509	TTATATTCTTTGGACCAAGATGGAGAGCTCTATACAGTCAGCAAAACCAAGACCAAGG	42568
QY	1230	AGCTTACTGTGGCTCAGATCATGAATCTCTTATTGCCCCAAATTCAGATTTAAATGAAGAA	1289
DB	42569	AGCTTACTGTGGCTCAGATCATGAATCTCTTATTGCCCCAAATTCAGATTTAAATGAAGAA	42628
QY	1290	AGTAGGAAAAACCACTAGATCACTCAGGTAAAGACCTAAATCCCTTTATGATTTATAC	1349
DB	42629	AGTAGGAAAAACCACTAGATCACTCAGGTAAAGACCTAAATCCCTTTATGATTTATAC	42688
QY	1350	AGTGAAGTGAAGAAATAGATTTAAGGGCCTAGATCTGTATAGACAGAGTACCTAATGAAT	1409

Db 42689 ATGGGAAGTCAGAAATAGATTGAAGGCTTAGATCTGATAGATAGAGTGCCTGATGACT 42748
Qy 1410 ATGGACAGAGTTTCATGACATTTGTACAGGACAGAGGATCGAGACCATCCCATGAGAAA 1469
Db 42749 ATGGAATGAGTTTCGTGAAGTTGTACAGGACAGAGGAATCAAGACCATCCCATAGAAA 42808
Qy 1470 GAAATGCAGAAAGCAAAATGCGTCTCTGGGAGGCTTACAAATAGCTGTGAAAGAG 1529
Db 42809 GAAATGAAAGCAAAATGCGTCTCTGGGAGGCTTACAAATAGCTGTGAAAGAG 42868
Qy 1530 AGAAGTGAAGCAAGCAAGAAAGAAAGAAAGATATAAGCATCTGAATGCAGAGTTCCAAA- 1588
Db 42869 AGAAGTGAAGCAAGCAAGAAAGAAAGAAAGATATAAGCATCTGAATGCAGAGTTCCAAAG 42928
Qy 1589 ----- 1588
Db 42929 AATAGCAAGGAGAGATAGAAAGCTTCTTTCAGCAATCAATGCAAGAAATAGAGGAAA 42988
Qy 1589 ----- 1588
Db 42989 CAACAGATGGGAAGACTAGGATCTCTTCAGAAATCAGAGATACCAAGAACAT 43048
Qy 1589 ----- 1588
Db 43049 TCCTGCAAGATGAGCTCGATAGAGGACAGAAATGGTATGGACCTAACAGAGCAGAGA 43108
Qy 1589 ----- 1588
Db 43109 TATTAAGAGAGATGGCAAGATACACAGAGAAGACTATACAAAAGATCTTCATGACCC 43168
Qy 1589 ----- 1588
Db 43169 AGATAATCAGATCTGTGATCACTGACTAGACCAGACATCTCTGGAATGTGAAGTCAA 43228
Qy 1589 ----- 1588
Db 43229 GTGGGCTTAGAAAGCATCACTAGAACAAAGCTAGTGGAGGCGATGGAATCCAGTTGA 43288
Qy 1589 ----- 1588
Db 43289 GCTATTCCAAATCCTGAAAGATGATGCTGTGAAAGTCTGCACCTCAATATGCCAGCAAT 43348
Qy 1589 ----- 1588
Db 43349 TTGGAAGCACTCAGCAGTGGCCACAGGACTGGAAAGGTCAGTTTTCATTCCAAATCCCAA 43408
Qy 1589 ----- 1588
Db 43409 GAAAGGCAATGCCAAAGATGCTCAAACTACCACAAATGGCACTCATCTCAATGCTAG 43468
Qy 1589 ----- 1596
Db 43469 TAAAGTAAAGCTCAAAATTTCTCCAAGCCAGGCTTCAGCAATATGTGAAACCGTGAACCTCC 43528
Qy 1597 AGTTGTTCAAGCTGTTTAGAAAGTCAGAGGACAGAGACCAATTCGCAACATCCT 1656
Db 43529 AGATATTCAAGCTGTTTAGAAAGTCAGAGGACAGAGATCAAAATGTGCAACATCTG 43588
Qy 1657 CTGTATCATGGAAGAAAGCAAGAGAGTTCAGAAAGCAATCTATTCTGCTTTTATGACTA 1716
Db 43589 CTGGATCATTGAAAGCAAGAGAGTTCAGAGGAACATCTATTCTGTTTATGACTA 43648
Qy 1717 TGCAGAAAGCTTTGACTGTGGGGTCAATATAACTGTGGAATTTCTGAAAGGATGGG 1776
Db 43649 TGCCAAAGGCTTTGACTATGTGATCAAAATTAATCTGTGAAATTCCTGAAAGAGATGGG 43708
Qy 1777 AATACCAGACACCTGACCTGACTCTTGAAAGATTTGTATGAGGTCAGAGCAACAGT 1836
Db 43709 AATACCAGACACCTGACCTGCTCTTGAGAAATTTGTATGAGGTCAGAGCAACAT 43768
Qy 1837 TAGAAGTCGATGGAACACAGACTGTTTCCAAAGTAGGAAAGGAGTATGTCAAGGCTG 1896

Db 43769 TAGAAGTCGACACGG--AAACAGACTGGTTCAAATAGGAAAGGAGTTGCTCAAGGCTG 43825
Qy 1897 TATATTGTCAACCCGGCTGTTTAACTTCTATGACAG--ACATCATGAGAAAGCTTGGGCT 1955
Db 43825 TATATTGTCAACCTGTTTATTTAATATATGACAGATACATCATGAGAAAGCTTGGACT 43885
Qy 1956 GGAAGAGCACAAGCTGGAATCAAGATTGCGGGAGAAATAGCAATAACCTCAGATATGCG 2015
Db 43885 GGAAGAGCACAAGCTGGAATCAAGATTGCGGGAGAAATATCAATAACCTCAGACATGCG 43945
Qy 2015 AGATGATACCCCTTATGCGAGAAAGTGAAGAGAACTAAAGCCCTTGTGATGAAGT 2075
Db 43945 AGATGACACCCCTTATGCGAGAAAGTGAAGAGAAATTCAAAGCCCTTGTGTTGAAAGT 44005
Qy 2075 GAAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTTCAGAAAGCAAGATCATGCG 2135
Db 44005 GAAAGTGAAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTTCAGAAAGCAAGATCATGCG 44065
Qy 2136 ATCTGCTCCCACTCTCATGGAATAGATGGGAAACAGTGGAAACAGTGTCCAGACTT 2195
Db 44066 ATCTGCTCCCACTCTCATGGAATAGATGGGAAACAGTGGAAACAGGCTCAGACTT 44125
Qy 2196 TATTTTGGGGGCTCCAAATCACTGCAGATGCTGACTGCGAGCATGAATTAAGAA 2255
Db 44126 TATTTT--CTGGGCTCCAAATCACTGCAGATGCTGACTGCGAGCATGAATTAAGAA 44183
Qy 2256 CACTTACTCTTGGAGAAAGTTA--ACCAACCTAGATAGCATATTGAAAGCAGAGAC 2313
Db 44184 CACTTACTCTTGGAGAAAGTTATGACCAACCTAGATAGCATATTCAAAGCAGAGAC 44243
Qy 2314 ATTACTTGCACAAAGCCCATCTAGTCAAGGCTATGGTTTTTCCAGTGTGATGAT 2373
Db 44244 ATTACTTGCACAACTAAGGTTCTGCTAGTCAAGGCTATGGTTTTTCTGTGCTATGAT 44303
Qy 2374 GGATGTGAGAGTTGAGCTGTGAAGAAAGCTGAGCACTGAAGAAATGATGCTTTTCAACTG 2433
Db 44304 GGATGTGAGAGTTGAGCCGTGAAGAGGCTGAGTCCGAGAAATGATGCTTTTCAACTG 44363
Qy 2434 TGGTGTGAGAGAGCTCTTGTGAGAGTCCCTTGGACTGCAAGAGAGATCCAAACAGTCCATT 2493
Db 44364 TGGTGTGAGAGAGCTCTTGTGAGAGTCCCTTGGACTGCAAGAGAGATCCAAACCAATCCATT 44423
Qy 2494 CTGAAGGAGATCAGCCCTGGGATTTCTTGGAGGAATGATGCTAAAGCTGAACTCCAG 2553
Db 44424 GTGAAGGAGATCAGCCCTGGGATTTCTTGGAGGAATGATGCTTAAAGCTGAACTCCAG 44483
Qy 2554 TACTTTGGCCACTGATCAGAAAGAGCTGACTCACTGGAAGAAAGACCTGATGCTGGAGGG 2613
Db 44484 TACTTTGGCCACTCATGCGAAGAGTTGATTCTATTGGGAAGAGACTGATACTGGAGGG 44543
Qy 2614 ATTGGGGCAGAGAGAGAGGAGCAGCAGAGGATGAGATGGCTGATGGCATCACTGAC 2673
Db 44544 GTTGGGGCAGAGAGTGAAGGGGACGACAGAGGATGAGATGGCTGATGGCAACACTGAC 44603
Qy 2674 TCGATGAGCTGAGTCTGGGTGAACTCCTGGAAGTTGGTGAAGCAGAGGAGGCTGTCTCT 2733
Db 44604 TCGATGAGCTGAGTCTGGGTGAACTCCTGGAAGTTGGTGAAGCAGAGGAGGCTTGGCT 44663
Qy 2734 GCGGCGAATTCATGGGGTCAAAAGAGTTGGACACGACTGAGCAACTGAACTGAACTGAAC 2793
Db 44664 GCTGCGAATTCATGGGGTTCMAAGAGTGGACACGACTGAGCAACTGATCTGATCTGATC 44723
Qy 2794 TGTACTGA 2801
Db 44724 TGTACTGA 44731

RESULT 37

AC108896/c

LOCUS

DEFINITION

ACCESSION

AC108896
Bos taurus clone RP42-89G3, WORKING DRAFT SEQUENCE, 5 ordered
pieces.
AC108896

223717 bp DNA linear
HTG 20-JUL-2002

VERSION AC108896.2 GI:21913015
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 223717)
 Akter, N., Antonellis, A., Ayle, K., Beckstrom-Sternberg, S.M.,
 Benjamin, S., Blackesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
 Brooks, S., Dietrich, N.D., Granite, S., Guan, X., Gupta, J.,
 Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P.,
 Lee, L.-S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
 Margulies, E.H., Masello, C., Maskeri, B., Mastrian, S.D.,
 McCloskey, J.C., McDowell, J., Paquirigan, C., Pearson, R.,
 Portnoy, M.E., Prasad, A., Schueler, M.G., Stantripop, S., Thomas, J.W.,
 Thomas, P.J., Touchman, J.W., Taurgeon, C., Vogt, J.D., Walker, M.A.,
 Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 223717)
 Green, E.D.
 Direct Submission
 Submitted (01-FEB-2002) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 223717)
 Green, E.D.
 Direct Submission
 Submitted (20-JUL-2002) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 On Jul 20, 2002 this sequence version replaced gi:18464072.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoo@hghri.nih.gov
 ----- Project Information
 Center project name: axh
 Center clone name: 089G03

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 222996 bases at least Q40
 Consensus quality: 223229 bases at least Q30
 Consensus quality: 223306 bases at least Q20
 Insert size: 198000; agarose-ff
 Insert size: 223317; sum-of-contigs
 Quality coverage: 12.57x in Q20 bases; agarose-ff
 Quality coverage: 11.14x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

* 1	3266:	contig of 3266 bp in length	
* 3267	3366:	gap of unknown length	
* 3367	53961:	contig of 50595 bp in length	
* 53962	54061:	gap of unknown length	
* 54062	115545:	contig of 61484 bp in length	
* 115546	115645:	gap of unknown length	
* 115646	200320:	contig of 84675 bp in length	
* 200321	200420:	gap of unknown length	
* 200421	223717:	contig of 23297 bp in length.	

FEATURES
 Location/Qualifiers
 1..223717
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-89G3"
 /clone_lib="RP42"
 1..111524
 /note="clone overlaps with GenBank Accession Number AC108176 clone RP42-57H8 (center project name axg)"
 1..21089
 /note="clone overlaps with GenBank Accession Number AC108175 clone RP42-250120 (center project name axf)"
 1..3266
 /note="assembly_fragment
 clone_end:77
 vector_side:left"
 3367..53961
 /note="assembly_fragment"
 54062..115545
 /note="assembly_fragment"
 115646..200320
 /note="assembly_fragment"
 200421..223717
 /note="assembly_fragment
 clone_end:SP6
 vector_side:right"

Query Match	52.9%;	Score 1890.91;	DB 1;	Length 223717;
Best Local Similarity	73.4%;	Pred. No. 0;		
Matches 2527;	Conservative	0;	Mismatches 291;	Indels 623; Gaps 15;

QY	1	GTCAGTAGGGGGCGAGTGAGGAGGAGTACCTACCTCGTCCAGGTAAAGGAGAGTACG	60
Db	92033	GTCAGTAGGAGCGCGCGTGAGGAGA-----TACCCCATGTCCAGGTAAAGGAGAGCGGC	91979
QY	61	TGCGTTTGTGGAGCAGCGGTAAAGAGATACCCACGCCCAAGGTAAAGAAACCCCAAG	120
Db	91978	TGTGCTTTGCTGGAGCAGCGGTGAAGAGATACCCCAAGTCCCAAGTAAAGAAACCCCAAG	91919
QY	121	TAAGATGTTAGTGTGTGTGAGAGGGGCATCAGAGGSCAGACATCTGAAACCATACACCA	180
Db	91918	TAAGATGTTAGTGTGTGTGAGAGGGGATCAGGCGCAGACAGATGAACCATACATCA	91859
QY	181	GAAGAACTAGTCAATCTAATCACTAGGACCAAGCCCTTGTTACTCAATCAATGAACCTAAG	240
Db	91858	GACAACTACCAATCTGTATCACAC--AGACACACAGCCCTTGTTACTCAATGAACCTAGG	91801
QY	241	CCATGCCGTGGGGCAACCCAGATGGGCGAGTCACTGGTGAGAGATCTGACAGATGTG	300
Db	91800	CCATGCCGTGGGGCCATCCCAAGCGGGGTGATGTGTGGAG--GGTCTGACAGATGTG	91742
QY	301	GTCCACTGGAGAAAGGGAAT-GCAAAACCACTTCAGTATTCTTGGCTTGAGAACCCCATGAA	359
Db	91741	GTCCACTGGAGAAAGGGAATGCAACCACTTCAGTATTCTTGGCTTGAGAACCCCATGAA	91682
QY	360	CAGTATGAAGGCAAAATGATAGATAGTGAAGAGGAATCCCCAGGTAGGTGTC	419
Db	91681	CAGTATGAAGGTTAAAGAGCAGGACGCTGAAAGATGAATCCCCAGGTGCGTAGGTGC	91622
QY	420	CCCATATGCTACTGGAGATCACTGTGAGAAATAACTCCAGAAAGATGAAGAGATGGAGCC	479
Db	91621	CCATATGCTACTGGAGATCACTGTGAGAAACAATCCAGAAAG-ATGAAGGATGAAGCT	91563

[illegible]

Db	133049	AAGCTAGTGGAGGTGATGGAATTCACAGTTCAGCTATTTCAAATCCTGAAAGATGATGCTG	132990
Qy	1589	-----	1588
Db	132989	TGAAATATTTCAGTCAATATGCGACAAATTTGGAAACTCAGCAGTGGCCACAGGACT	132930
Qy	1589	-----	1588
Db	132929	GGAAAGGTACCTTTTCGTTCCAAATCCCAAGAAAGGCATGTCAAGAATGCTCAAACT	132870
Qy	1589	-----	1588
Db	132869	ACTGCACAATTGCACCTCATCTCACACGCTAGTAAAGTAAGCTCAAAATTTCTCCAAGCCA	132810
Qy	1589	-----GAACTTCCAGTTGTTCAAGCTGGTTTATAGAAAAGTCA	1625
Db	132809	GGCTTCAGCAATACGTGAACCTGGAACCTTCAGATGTTCAAGCTGGTTTATAGAAAAGGCA	132750
Qy	1626	GAGGAACAGAGACCAAAATTCGCAACATCCTCTGTATCATGGAAGAAAGCAAGAGATTCC	1685
Db	132749	GAGCAACAAGAGATTAAATTTGCCAATCCGCTGGATCATGGAAGAAAGCAAGAGATTCC	132690
Qy	1686	AGAAACATCATTTCTGCTTTATTTGACTATGCAAGAGCTTTGACTGTGGGGTCA	1745
Db	132689	AGAAACATCATTTCTGCTTTATTTGACTATGCAAGAGCTTTGACTGTGGGATCAGA	132630
Qy	1746	ATAAACTGTGGAATTTCTGAAAGGATGGGAATPACCAGACCACTGACCTGCTTTGA	1805
Db	132629	ATAAATGTGGAATTTCTCAAGAGTTGGGAATPACCATTCATCTGACCTGCTTTGA	132570
Qy	1806	AAAATTTGTATGAGGTGAGGAACACAGTTAGAACTGACATGGAACAACAGACTGGT	1865
Db	132569	TGAATCTGTATGCGAGTCAGGAACACAGTTAGAACTGGAACCAAAAGACTGGC	132510
Qy	1866	TCCAAGTAGGAAAAGGAGTAGTCAAGCTGTATATTGTCAACCGGCTTGTAACTTCT	1925
Db	132509	TCTAAATTTGGAAAGGAGTAGCTCAAGCTGTATATTGTCAACCTGCTTATTAACTTAT	132450
Qy	1926	ATGCAGAG-ACATCATGAGAAACCTGGCTGGGAAGAACACAGCTGGGAATCAAGATTG	1984
Db	132449	CTGCAGAGCACATCATGAGAAA--CTGGGCTGGAGGAAGCACAGCTGGGAATCAAGACTG	132392
Qy	1985	CCGGGGAATAGCAATAAATCACTCAGATATGAGATATGAGATATACACCTTTATGCGAAGAGTG	2044
Db	132391	CCTAGAGAAATATCAGTAATCTCAGATACGAGATGACACACCTTTATTGCGAAGGTA	132332
Qy	2045	AAGAGGAATAAAAGGCTCTTTGATGAAGGTGAAAGAGAGAGTGAAAAAGTTGGCTTAA	2104
Db	132331	AAGAAGAACTAAGAGGCTCTTTGATGAAGGAAGGAAGGAAGTAAAAAAAGTTGGCTGAA	132272
Qy	2105	AGCTCAACATTCAGAAACAGAGATCATGCACTCTGCTCCATCCTCATCTGAGGAATAG	2164
Db	132271	AGCTCAATATTTCAGAAAATGAGATCATGCGATCTGCTCCATCCTCATCTGAGGAATAG	132212
Qy	2165	ATGGGAAACAGTGAAACAGTGTCAAGCTTTATTTTGGGGGCTCCAAAATCACTGC	2224
Db	132211	ATGGGAAATAGTCGAACAGTGGCTGACTTTATTTCT--TGGGCTCCAAAATCACTGC	132154
Qy	2225	AGATGGTGAATGCGACCAATGAAATTAAGACACTTACTCTCTGGAGAAAGTTA--AC	2282
Db	132153	AGATGGTGAATGCGACCAATGAAATTAAGACACTTACTCTCTGGAGAAAGTTATGAT	132094
Qy	2283	CAACTAGATAGCATTTGAAAGCAGAGACATTAACCTTGGCAACAAGGCCCATCTAGT	2342
Db	132093	CAACTAGATAGCATTTGAAAGCAGAGACATTAACCTTGGCAACAAGGCCCATCTAGT	132034
Qy	2343	CAAGCTATGTTTTTCCAGTGGTCAATGATGATGATGAGATTGGAGTTGGAGTGAAGCAAGC	2402
Db	132033	CAAGCTATGTTTTTCCAGTGGTCAATGATGATGATGATGAGATTGGAGTTGAAGCAAGC	131974
Qy	2403	TGAGCACTGAAGATTGATGCTTTTGAACCTGTTGTTGGAGAGACTCTTGAGAGTCCC	2462
Db	131973	TCAGCACAAAATAATGATGCTTTTGAACCTGTTGTTGGAGAGACTCTTGAGAGTCTC	131914
Qy	2463	TTGGAATCAAGGAGATCCAAACCACTGCTCAATTTCTGAAGGAGATCAACCTCGGATTTCTTT	2522
Db	131913	ATGGACTACAGGAGATCCAAACCACTGCTCAATTTCTGAAGGAGATCAGTCTCGGTTCTTT	131854
Qy	2523	GGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTGGGCCACCTGATCAGAGAGCTGA	2582
Db	131853	GGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTGGGCCACCTGATCAGAGAGTTGA	131794
Qy	2583	CTCAGTCAAAAGACCTCTGATGCTGGAGGAGTATGGGGGAGGAGGAGGAGGAGGAGCA	2642
Db	131793	CTCATGTGAAAAGATCCTGATTTCTGGGAAAGATTTGAGTAGGAGAGTAGGGATGACA	131734
Qy	2643	GAGGATGAGATGGCTGGATGGCATCAGTCTGATGGACCTGAGTCTGGGTGAACCTCT	2702
Db	131733	GAGGATGAGATGGCTGGATGGCATCAGTCTGATGGACCTGAGTCTGGGTGAACCTCT	131674
Qy	2703	GGAGTTGCTGATGACAGGAGGAGGCTGCTCGCGCGATTCATGGGGTCACAAAGAGTTG	2762
Db	131673	GGAGTTGCTGATGACAGGAGGAGGCTGCTCGCGCGATTCATGGGGTCACAAAGAGTTG	131615
Qy	2763	GACAGACTGAGCAACTGAACTGAACTGTAAGTGAACCTTAGTAACTTTATATTAC	2822
Db	131614	GACATGCTGAGCAACTGAACTGAACTGTAAGTGAACCTTAGTAACTTTATATTAC	131555
Qy	2823	TCAGAAAATAGTAATTTTATA 2843	
Db	131554	TCTGGAGAATTAATGACAGA 131534	
RESULT 39	AC137647/c	165992 bp DNA linear HTG 04-FEB-2003	
LOCUS	AC137647.2	GI:28201542	
DEFINITION	Bos taurus clone RP42-127F18, WORKING DRAFT SEQUENCE, 6 ordered		
ACCESSION	AC137647		
VERSION	AC137647.2	GI:28201542	
KEYWORDS	HTG; HTGS; PHASE2; HTGS_DRAFT.		
SOURCE	Bos taurus (cow)		
ORGANISM	Bos taurus		
REFERENCE	1 (bases 1 to 165992)		
AUTHORS	Akhter N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Carliaga, K., Coleman, S., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.B., Maduro, V.B., Marulies, E.H., Masello, C., Maskeri, B., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sisor, C., Stantripo, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, J., Young, A. and Green, E.D.		
TITLE	NISC Comparative Sequencing Initiative		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 165992)		
AUTHORS	Green, E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-NOV-2002) NIH Intramural Sequencing Center, 8717		
REFERENCE	Groveton Circle, Gaithersburg, MD 20877, USA		
AUTHORS	Green, E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-FEB-2003) NIH Intramural Sequencing Center, 8717		
REFERENCE	Groveton Circle, Gaithersburg, MD 20877, USA		
COMMENT	On Feb 4, 2003 this sequence version replaced gi:25699231.		
	----- Genome Center		
	Center: NIH Intramural Sequencing Center		
	Center code: NISC		
	Web site: http://www.nisc.nih.gov		
	Contact: nisc_zoo@nhgri.nih.gov		

ACCESSION VERSION KEYWORDS SOURCE ORGANISM

AC129960
AC129960.2 GI:27901861
HTG; HTGS; PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 190441)
Bos taurus, clone RP42-362K14
2 (bases 1 to 190441)
Unpublished

REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, I., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hags, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL REFERENCE AUTHORS

Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 190441)

Birren, B., Nusbaum, C., Lander, E., Abouelheil, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hags, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL COMMENT

Submitted (25-JAN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 25, 2003 this sequence version replaced gi:22123282.
All repeats were identified using RepeatMasker:
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L24461

Center clone name: 362_K.14

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731
Consensus quality: 190436 bases at least Q40
Consensus quality: 190441 bases at least Q30
Consensus quality: 190441 bases at least Q20
Insert size: 198000; agarose-fp
Insert size: 190441; sum-of-contigs
Quality coverage: 16.1 in Q20 bases; agarose-fp
Quality coverage: 16.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 190441: contig of 190441 bp in length.
assembly_fragment_location/Qualifiers
1. 190441
/organism="Bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="RP42-362K14"
/clone_lib="RPC1-42 Bovine BAC"

FEATURES source

Query Match	Score	DB 1;	Length	DB 1;	Length	DB 1;	Length
62	GGCTTTCTGTCGAGCAGCGGTAAAGAGATACCCAGCCCGGATACCCAGGTAAGAGAACCCCAAGT	121					
137961	GGCTTTCTGTCGAGTAGCGGTAAAGAGATACCCAGGTAAGAGAACCCCAAGT	137902					
122	AAAGTGTAGTGTGTTGTGAGAGGGGATCAGAGGGGAGAGATCTGAAACCCATACACGAG	181					
137901	AAAGCAGTAGGCACTGAGAGAGGGGATCAGAGGGGAGAGATCTGAAACCCATACACGAG	137842					
182	AAACCTAGTCACTTAATCACTAGGACCAAGCTTGTCTAATCTCAATGAATAGC	241					
137841	ACAACAATCTGAT-----CACATGGACCAAGCTTGTCTAATGAATAGC	137789					
242	CATGCCGTGGGGCAACCCAGATGGCAGGTCATGTTGGAGAGATCTGACAGATGTGG	301					
137788	GGGGCCACCCAG--ACCCACATGTTGGGTGATGTTGGAGAGTCTGACAGATGTGG	137731					
302	TCCATGGAGAGGGAAT-GCAAAACCACTTCAGTATTTTGCCTTGAGAACCCCATGAAC	360					
137730	TCCACTGAAGAGGGAATGGCGAACCACTTCAGTATTTTGCCTTGAGAACCCCATGAAC	137671					
361	AGTATGAAAGGCAAAATAGTAGGATCTGAAAGAGGAATCCCGAGGTGAGTGGCC	420					
137670	AGTATGAAAGGCAAAATAGTAGGATCTGAAAGAGGAATCCCGAGGTGAGTGGCC	137611					
421	CCATATGCTTCTGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGAGATGAGGCA	480					
137610	CAATATGCTTCTGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGAGATGAGGCA	137551					
481	AGGCAAAAGATACCCAGCTGTGGATGTGATGTGATTAAGCAAGGTGCGATGCTGT	540					
137550	AGGCAAAAGATACCCAGCTGTGGATGTGATGTGATTAAGCAAGGTGCGATGCTGT	137491					
541	AAAGAGCAATATTGCAATAGGAACCTGGAATGTTCAGGTCCATGAATCAAGGCAAAATGGAA	600					
137490	AAAGAGCAATATTGCAATAGGAACCTGGAATGTTCAGGTCCATGAATCAAGGCAAAATGGAA	137431					
601	GTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAGGCAATAAAA	660					
137430	GTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAATCAGTGAATAAAA	137371					
661	TGGACTGGNAATGGGTGAATTTAACTCAGATGACCATTTATCTACTCTGCGGGCAGAA	720					

Db 61388 ATGGGAATACAGACCATCTGCTCTGCTCTTGAGAAACCTATATGACGGTCAGGAAGCA 61447
 QY 1832 ACAGTTAGAACTGGACATGGAACACAGACATGGTTCCAGTAGGAAAAAGGAGTATGTCAA 1891
 Db 61448 AAGATTAACTGGATATGGAACACAGACATGGTTCCAAATAGGAAAAAGGAGTACATCAA 61507
 QY 1892 GCCTGTATATTGTCACCCGGCTGTTTAACTTCTATGACAGAG-ACATCATGAGAAACGCT 1950
 Db 61508 GCCTGTATATTGTCACCTGCTTATTAACTTATATGACAGATACATCATGAGAAACGCT 61567
 QY 1951 CGGCTGGAAGACACAAAGCTGGAAATCAAGATTGCCGGGAGAAATAGCAATAAATCACTCAGA 2010
 Db 61568 CGGTTGGAAGACACAAAGCTGGAAATCAAGATTGCCGGGAGAAATATCAA-AAACTCAGA 61626
 QY 2011 TATGACAGATATACACACCTTATGCGACAAAGTGAAGGAACTAAAAAGCCTCTTCTATG 2070
 Db 61627 TATGACAGATATACACACCTTATGCGACAAAGTGAAGGAACTAAAAAGCCTCTTCTATG 61686
 QY 2071 AAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAGATC 2130
 Db 61687 AAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAGATC 61746
 QY 2131 ATGCACTGCTGCTCCATCACTTCATGGAATAGATGGGAAACAGTGGAAACAGTGTCA 2190
 Db 61747 ATGCACTGCTGCTCCATCACTTCATGGAATAGATGGGAAACAGTGGAAACAGTGTCA 61797
 QY 2191 GACTTTATTTTGGGGGCTCCAAATCACTGCAGATGCTGACGAGCATGAAATTA 2250
 Db 61798 GACTTTATTTT--CTGGGCTCCAAATCACTGCAGATGCTGACGAGCATGAAATTA 61855
 QY 2251 AAAGACACTTACTCTTGGAAAGAAAGTTA--ACCACTAGATAGCATATTTGAAAGCA 2308
 Db 61856 AAAGACGCTTACTCTTGGAAAGAAAGTTATGACCAACCTAGATAGCATATTTCAAAGCA 61915
 QY 2309 GAGACATTACCTTGCACAAAGCCCATCTAGTCAAGGCTATGTTTTTCCAGTGTCTCA 2368
 Db 61916 GAGACA-----TTGCCACAAAGTCCATCTAGTCAAGGCTATGTTTTTCCAGTGTCTCA 61970
 QY 2369 TGTATGATGAGAGTTGGATCTGTGAAGAAAGCTGACCACTGAAGAAATGATGCTTTTG 2428
 Db 61971 TGTATGATGAGAGTTGGATCTGTGAAGAAAGCTGACCGCTGAAGAAATGATGCTTTTG 62030
 QY 2429 AACTGTGTGTTGGAGAGACTCTTGAGAGTCCCTTGGATCGCAAGAGATCCAAACAGT 2488
 Db 62031 AACTGTGTGTTGGAGAGACTCTTGAGAGTCCCTTGGATCGCAAGAGATCCAAACAGT 62090
 QY 2489 CCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCTAAAGCTGAAC 2548
 Db 62091 CCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCTAAAGCTGAAC 62150
 QY 2549 TCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTGCTGGAAGAAAGCCTGATGCTGG 2608
 Db 62151 TCCAGTACTTTGGCCACCTGATGTAAGAGTTGACTCATTTGGAAGAAAGCCTGATGCTGG 62210
 QY 2609 GAGGATTTGGGGCAGGAGAGAGAGGACACAGAGATGAGATGCTGATGGCATCA 2668
 Db 62211 GAGGATTTGGGGCAGGAGAGAGAGGACACAGAGATGAGATGCTGATGGCATCA 62270
 QY 2669 CTGACTCGATGAGAGTGTGGGTGAATCTCTGGAATCTGATGATGATGATGATGATGATG 2728
 Db 62271 CTGA--TCGATGAGATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATG 62329
 QY 2729 GTCTGCGGGCATTCATGGGCTCAAGAGATTGGACACGACTGAGCAACTGACTGAAC 2788
 Db 62330 GGCCTGCTGCCATTCTGGGGTCAAAAGAGTCCGACACGACTGAGCAACTGACTGAAC 62389
 QY 2789 TGA 2791
 Db 62390 TGA 62392

LOCUS AC109787 154577 bp DNA linear HTG 05-DEC-2002
 DEFINITION Bos taurus clone RP42-136015, WORKING DRAFT SEQUENCE, 3 ordered
 pieces.
 AC109787
 VERSION AC109787.2 GI:26050954
 KEYWORDS HTG; HTGS PHASE2; HTGS_DRAFT.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 154577)

AUTHORS Benjamini, B., Blakesley, A., Ayele, K., Beckstrom-Sternberg, S.M.,
 Carlaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
 Raghathi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
 Larcic, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, O.L., Maduro, V.B.,
 Margulies, E.H., Masello, C., Maskeri, B., McDowell, J.,
 Parguligan, C., Pearson, R., Portnoy, M.E., Prasad, A.,
 Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,
 Stantridop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
 Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 154577)
 Green, E.D.

TITLE Direct Submission
 JOURNAL
 REFERENCE 2 (bases 1 to 154577)
 AUTHORS Green, E.D.

TITLE Submitted (07-FEB-2002) NIH Intramural Sequencing Center, 8717
 JOURNAL Grovemont Circle, Gaithersburg, MD 20877, USA
 REFERENCE 3 (bases 1 to 154577)
 AUTHORS Green, E.D.

TITLE Direct Submission
 JOURNAL Submitted (05-DEC-2002) NIH Intramural Sequencing Center, 8717
 REFERENCE 4 (bases 1 to 154577)
 AUTHORS Grovemont Circle, Gaithersburg, MD 20877, USA
 COMMENT On Dec 5, 2002 this sequence version replaced gi:18581925.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Contact site: http://www.nisc.nih.gov
 Web site: nisc.zoo@hghri.nih.gov
 ----- Project Information
 Center project name: cjr
 Center clone name: 136015

The sequence data in this record represents an 'enhanced'
 version of a Phase 2 submission. Specifically, the indicated
 order and orientation of each sequence contig has been
 established using one or more of the following: read-pair
 data from individual subclones, overlaps with neighboring
 clones, alignment with available reference sequence (e.g.,
 human), and/or confirmation by PCR testing. In addition,
 the sequence assembly is based on at least 8x average
 coverage in Q20 bases and has been reviewed to rule out
 gross misassemblies, the low-quality ends of sequence
 contigs have been trimmed away, and each base is associated
 with a Phrap-derived quality score.

----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 154307 bases at least Q40
 Consensus quality: 154347 bases at least Q30
 Consensus quality: 154358 bases at least Q20
 Insert size: 121000; agarose-fp
 Quality coverage: 12.74x in Q20 bases; agarose-fp
 Quality coverage: 9.99x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have

* provided by the submittor.

* This sequence will be replaced

- * by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 88559: contig of 88559 bp in length

* 88560 88659: gap of unknown length

* 88660 126656: contig of 37997 bp i

*	126657	126656: concy	of unknown length
*	126657	126656: concy	of unknown length

* 126757 contig of 27821 bp in length.
126758: gap of unknown length

FEATURES

source

T. .154511
/orajom=

```
/organism="Bos taurus"
```

```
/mol_type="genomic DNA"
```

/db_xref="taxon:9913"

```
/clone="RP42-136O15"
```

```
/clone_lib="RP42"
```

1. .88559

```
I.:00000  
/note="assembly fragment
```

```
clone end:sp6
/note="assembly_fragment"
```

```
clone_end:SP6
vector_size:left"
```

```
vector_side:left"
1 1453
```

1. .14623

```
/note="clone overlaps wi
```


data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 199764 bases at least Q40
 Consensus quality: 200000 bases at least Q30
 Consensus quality: 200271 bases at least Q20
 Insert size: 157000; agarose-fp
 Insert size: 200718; sum-of-contigs
 Quality coverage: 12.42x in Q20 bases; agarose-fp
 Quality coverage: 9.72x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently:
 * consists of 5 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 70538: contig of 70538 bp in length
 * 70539 70638: gap of unknown length
 * 70639 77542: contig of 6904 bp in length
 * 77543 77642: gap of unknown length
 * 77643 133416: contig of 55774 bp in length
 * 133417 133516: gap of unknown length
 * 133517 149909: contig of 16393 bp in length
 * 149910 150009: gap of unknown length
 * 150010 201118: contig of 51109 bp in length.

FEATURES

source

1. .201118
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9913"
 /clone="RP42-71E11"
 /clone_lib="RP42"

misc_feature

1. .98740
 /note="clone overlaps with GenBank Accession Number
 AC109787 clone RP42-136015 (center project name cjr)"

misc_feature

1. .70538
 /note="assembly_fragment
 clone end.SP6
 vector_side:left"

misc_feature

70639. .77542
 /note="assembly_fragment"

misc_feature

77643. .133416
 /note="assembly_fragment"

misc_feature

133517. .149909
 /note="assembly_fragment"

misc_feature

150010. .201118
 /note="assembly_fragment
 clone_end:T7
 vector_side:right"

misc_feature

153492. .201118
 /note="clone overlaps with GenBank Accession Number
 AC109797 clone RP42-58E2 (center project name cjp)"

Query Match

Best Local Similarity 52.2%; Score 1864.4; DB 1; Length 201118;

Matches 2288; Conservative 0; Mismatches 126; Indels 596; Gaps 5;

Qy

380 ATAGGACTGAAGAGAACTCCCGAGGTGAGTGGCCCATGTCTACTGGAGATC 439

Db

30436 ATAGGACTGAAGAGAACTCCCGAGGTGAGTGGCCCATGTCTACTGGAGATC 30377

Qy	440	AGTGGAGAAATAACTCCAGAAAAGAAATGAAGAGATGAGCCAAAGCAAAAAGAAATACCCAG	499
Db	30376	AGTGGAGAAATAACTCCAGAAAAGAAATGAAGAGATGAGCCAAAGCAAAAAGAAATACCCAG	30317
Qy	500	CTGTGGATGTGATCTGGTGTATATAAGCAAGGTCCGATGCTGTAAAGAGCAATATTGTCATAG	559
Db	30316	CTGTGGATGTGATCTGGTGTATATAAGCAAGGTCTGTATGCTGTAAAGAGCAATATTGTCATAG	30257
Qy	560	GAACCTGGATCTCAGGTCCATGAATCAAGCAAAATGGAGTGGTCAACAGAGATGG	619
Db	30256	GAACCTGGATCTCAGGTCCATGAATCAAGCAAAATGGAGTGGTCAACAGAGATGG	30197
Qy	620	CAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAACTAAATGGACTGGATGGTGAAT	679
Db	30196	CAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAACTGAATGGACTGGATGGTGAAT	30137
Qy	680	TTAACTCAGATGACCATTTATATCTACTACTTGGGGCAGGAATCCCTCAGAGAAATGGAG	739
Db	30136	TTAACTCAGATGACCATTTATATCTACTACTTGGGGCAGGAATCCCTCAGAGAAATGGAG	30077
Qy	740	TAGCCATCATGTCAACAAAAGAGTCCGAAATGCGAGTACTTGGATGCGAGTCTCAAAAACG	799
Db	30076	TGCCCCATCAGGTCAACAAAAGAGTCCGAAATGCGAGTACTTGGATGCGAGTCTCAAAAACG	30017
Qy	800	ACAGAATGATCTCTGTTTGTTCCTCAAGGCAACCAATTCATATACAGTATATCCAAAGTCT	859
Db	30016	ACAGAATGATCTCTGTTTGTTCCTCAAGGCAACCAATTCATATACAGTATATCCAAAGTCT	29957
Qy	860	ATGCCCCACACAGTAACTGCTGAAGAGCTGAAGTTGAACGGTCCCTATGAAGACCTTACAAAG	919
Db	29956	ATGCCCCACACAGTAACTGCTGAAGAGCTGAAGTTGAACGGTCCCTATGAAGACCTTACAAAG	29897
Qy	920	ACCTTTTAGAATCAACACCCCAAAAAGAGTCTCTTCTCATATATAGGGGACTGGAATGAA	979
Db	29896	ACCTTTTAGAATCAACACCCCAAAAAGAGTCTCTTCTCATATATAGGGGACTGGAATGAA	29837
Qy	980	AGTAGGAGCAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCTTTGGAATACGGAGT	1039
Db	29836	AGTAGGAGTCAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCTTTGGAATACGGAGT	29777
Qy	1040	AAGCAGGGCAAGAGATAATAGAGTTTTTCCAGAAAGATGCACTGGTCTATAGCAACACCC	1099
Db	29776	AAGCAGGGCAAGAGATAATAGAGTTTTTCCAGAAAGATGCACTGGTCTATAGCAACACCC	29717
Qy	1100	TCTTCCAAACACAGAGAGAGCTCTACACATGGAGATCAACAGATGGTCAACACCCAA	1159
Db	29716	TCTTCCAAACACAGAGAGAGCTCTACACATGGAGATCAACAGATGGTCAACACCCAA	29657
Qy	1160	ATCAGATTGATTATATCTTTTCAGCCAAAGATGGAGAGCTCTATACAGTCAGCAAAA	1219
Db	29656	ATCAGATTGATTATATCTTTTCAGCCAAAGATGGAGAGCTCTATACAGTCAGCAAAA	29597
Qy	1220	CAAGACAGGAGCTTACTGTGGCTCAGATCATGAACTCCTTATTGCCAAATTCAGACTTA	1279
Db	29596	CAAGACAGGAGCTTACTGTGGCTCAGATCATGAACTCCTTATTGCCAAATTCAGACTTA	29537
Qy	1280	AATTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAGAGCTTAATCCATCCCTT	1339
Db	29536	AATTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAGAGCTTAATCCATCCCTT	29477
Qy	1340	ATGATTATACAGTGGAGAAATAGATTAAAGGGCTTAGATCTGTATAGACAGATGAC	1399
Db	29476	ATGATTATACAGTGGAGAAATAGATTAAAGGGCTTAGATCTGTATAGACAGATGAC	29417
Qy	1400	CTAATGAATATGGACAGAGTTCAATGACATGTACAGAGACAGGATCCAGACCATCC	1459
Db	29416	CTAATGAATATGGACAGAGTTCAATGACATGTACAGAGACAGGATCCAGACCATCC	29357
Qy	1460	CCATGGAAAAAGAAATGCAAAAAAGCAAAATGGCTGTCTGGGGAGGCCCTTACAAATAGCTG	1519
Db	29356	CCATGG-AAAAGAAATGCAAAAAAGCAAAATGGCTGTCTGGGGAGGCCCTTACAAATAGCTG	29298

Qy	1520	TGAAAAGAGAGAAAGTGA	AAAAAGCAAAAGGAAAAA	AGGAAAAAGATAAAAGCATCT	GAATGCAG	1579
Db	29297	TGAAAAGAGAGAAAGCA	AAAAAGCAAAAGGAAAAA	AGGAAAAAGATAAAAGCATCT	GAATGCAG	29238
Qy	1580	AGTTCCAAA-----				1588
Db	29237	AGTTCCAAAGATAGCA	AGAGAGATAGAAAGCC	TTCTCAGTGATCAATG	CAATGCAAAAGAAA	29178
Qy	1589	-----				1588
Db	29177	TAGAGAAAAACAAGAT	AGGAAAGACTAGAGAT	CTCTTCAAGAAAAAT	CAGAGATACCA	29118
Qy	1589	-----				1588
Db	29117	AAGGAACATTTCAAT	GCAAGATGGCGTCAAT	TAAGACAGAAATGG	TATGGACCTTAACAG	29058
Qy	1589	-----				1588
Db	29057	AAGCAGAAGATATTA	AGAGAGAGATGCAAG	AANTACACAGAGAA	CTGTACAAAAAAGATC	28998
Qy	1589	-----				1588
Db	29997	TTCAATGACCAGATA	TAATCAGATGGTGT	GATCACTGACCTAG	ACGACACATPCCTGG	28938
Qy	1589	-----				1588
Db	28937	GTGAAGTCAAGTGG	CCCTTAGAATCACT	ATGGACAAGCTAG	TGGAGGTGATGGAATTC	28878
Qy	1589	-----				1588
Db	28877	AGTTGAGCTATTT	CAAAATCCTAAAG	ATGATGCTGTGAAG	TGCTGCACTCAATATG	28818
Qy	1589	-----				1588
Db	28817	GCAAAATTG	GAANAATCAGCAG	TGGCCACAGACTG	GAANAAGTCAGTTTC	28758
Qy	1589	-----				1588
Db	28757	CCCAAGAAAGGCA	ATGTCAAAGATGCT	CAAACTACCGCA	CAATGTCACATCTCACA	28698
Qy	1589	-----				1590
Db	28697	CGCTAGTAAGT	TAATGCTTAAANA	ATTCCTCAAGCCAG	GGCTTCAGCAATACTTGA	28638
Qy	1591	ACTTCCAGTTGTT	CAAGCTGGTTT	TAGAAAAGTCA	GAGGAACGACAGACCA	1650
Db	28637	ACTTCCAGTTGTT	GAAAGCTGGTTT	TAGAAAAGGCA	CGAACGAGATCA	28578
Qy	1651	CATCCTCTGTAT	CATGGAANAAC	AGAGAGTTC	CAGAAAAACATCTATT	1710
Db	28577	CATATGCTGGAT	CATGGAANAAC	AGAGAGTTC	CAGAAAAACATCTATT	28518
Qy	1711	TGACTATGCA	AAAAAGCCCTTT	GACTGTGGGGT	CACAATAAATCTG	1770
Db	28517	TGACTATGCC	AAAGCCCTTT	GACTGTGGTGG	ATCAATNAATCTG	28458
Qy	1771	GATGGGAATAC	CAAGCACTGT	ACCTGACTCT	TTGAAAAATTTG	1830
Db	28457	GATGGGAATAC	CTGACCTCTG	ACCTCTTTG	AGAAATCTGTATG	28398
Qy	1831	AAACAGTTAG	AACTGGACATG	GAACAACAGACT	GTGTTCCAAAGTAG	1890
Db	28397	AACATTAG	AACTGGACATG	GAACAACAGACT	GTATTC	28338
Qy	1891	AGGCTGTAT	TGTCACCGGCT	TGTTAACTCT	TATGCAGAG-ACATCAT	1949
Db	28337	AGGCTGTAT	TGTCACCGCT	CTCTTAACTCT	TATGCAGAGTACATCAT	28278
Qy	1950	TGGGCTG	GAAGACCA	CRAGCTGG	AAATCAAGATTGCC	2009
Db	28277	TGGGCTG	GAAGACCA	CRAGCTGG	AAATCAAGATTGCC	28218
Qy	2010	ATATGCAG	ATGATACCA	CCCTTATG	CGCAAAAGTGAAG	2069

Db	28217	ATATGCAGATGACACACACCTTATATGGCAGAAAAGTGAAGAGGAACCTAAAAAAGCCTCTTGAT	28158
Qy	2070	GAAGGTTGAAAGAGGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAAGAAAT	2129
Db	28157	GAAGGTTGAAGAGGACAGTGAATGTTGGCTTAAATCTCAACATTCGAAAAAGAAAT	28098
Qy	2130	CATGGCATCTGTCCTCCATCACCTCATGGAATAGATGGGAAACAGTGGAAAAACAGTGTCT	2189
Db	28097	CATGGCATCTGTCCTCCATCACCTCATGGAATAGATGGGAAACAGTGGAAAAACAGTGTCT	28038
Qy	2190	AGACTTTATTTTTTGGGGGGCTCCAAATCACTGCAGATGTTGACTGCAGCCATGAAAT	2249
Db	28037	AGACT--TATTTTGGGGGGCTCCAAATCACTGCAGATGTTGACTGCAGCCATGAAAT	27981
Qy	2250	AAAAGACATTACTCTCTGGAAAGAAAGTTA--ACCACCTAGATAGCATATTCAAAAGC	2307
Db	27980	AAAAGACGCTTACTCTCTGGAAAGAAAGTTATGACCAACCTAGATAGCATATTCAAAAGC	27921
Qy	2308	AGAGCATTAACCTTGGCCAAACAAAGCCCATCTAGTCAGGCTATGTTTTTCCAGTGGTC	2367
Db	27920	AGAGCATTAATTTTGGCCAAACAAAGGTCGTCTAGTCAGGCTATGTTTTTCCAGTGGTC	27861
Qy	2368	ATGTATGGATGTGAGAGTTGAGCTGTGAAGAAAGCTGAGCACTGAAGAAATTCATGCTTTT	2427
Db	27860	ATGTATGGATGTGAGAGTGGACTGTGAAGAAAGCTGAGCTGAAGAAATTCATGCTTTT	27801
Qy	2428	GAATCTGTTGTTGGAAAGACTCTTGAGAGTCCCTTGGACTGCAAGAGAGATCCAAACAG	2487
Db	27800	GAATCTGTTGTTGGAAAGACTCTTGAGAGTCCCTTGGACTGCAAGAGAGATCCAAACAG	27741
Qy	2488	TCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTTGAAGGAATGATGCTAAAGCTGAAA	2547
Db	27740	TCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTTGAAGGAATGATGCTAAAGCTGAAA	27681
Qy	2548	CTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCACTGGAAAAAGCCCTGATGCTG	2607
Db	27680	CTGTAGTACTTTTGGCCACCTGATCAGAGAGTGGACTCATTTGGAAAAAGACTCTGATGCTG	27621
Qy	2608	GGAGGATTTGGGGCAGGAGGAGAGGGGACGACAGAGATGAGATGCTGGATGGCATC	2667
Db	27620	GGAGGATTTGGGGCAGGAGGAGAGGGGATGACAGAGATGAGATGCTGGATGGCATC	27561
Qy	2668	ACTGACTTCGATGACGTGAGTCTGGTGAATCTCGGATTCGATGGATGGACAGGAGGCC	2727
Db	27560	ACTGACTTCGATGACGTGAGTCTGGTGAATCTCGGATTCGATGGATGGACAGGAGGCC	27501
Qy	2728	TGTCTCTGGCGGATTCATGGGCTCACAAGAGTTGGACACGACTGAGCAACTGAACTGAA	2787
Db	27500	TGGCGTGTCTGCAATTCATGGGCTCGCAAGAGTTCGACACGACTGAGTGTGAACCTGAA	27441
Qy	2788	CTGAACCTGTA 2797	
Db	27440	CTGAACCGAA 27431	
RESULT 44			
AC137653			
LOCUS			
DEFINITION	AC137653	121609 bp	DNA linear
	Bos taurus	clone RP42-506N12,	WORKING DRAFT SEQUENCE, 6 ordered
		pieces.	
ACCESSION	AC137653		
VERSION	AC137653.2	GI:282261477	
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.		
SOURCE	Bos taurus		
ORGANISM	Bos taurus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;		
	Bovidae; Bovinae; Bos.		
	1 (bases 1 to 121609);		
REFERENCE			
AUTHORS	Akhtar, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,		
	Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,		
	Carriaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,		

Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E.,
 Latic, P., Lee-Lin, S.-Q., Legaapi, R., Macuro, Q.L., Maduro, V.B.,
 Margulies, E.H., Mastello, C., Maskeri, B., McDowell, J.,
 Paquirigan, C., Pearson, R., Portnoy, M.E., Prasad, A.,
 Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C.,
 Startripp, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L.,
 Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.

TITLE JOURNAL

REFERENCE AUTHORS

TITLE JOURNAL

REFERENCE AUTHORS

TITLE JOURNAL

COMMENT

Unpublished
 NISC Comparative Sequencing Initiative
 2 (bases 1 to 121609)

Green, E.D.

Direct Submission

Submitted (27-NOV-2002) NIH Intramural Sequencing Center, 8717

Grovenmont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 121609)

Green, E.D.

Direct Submission

Submitted (06-FEB-2003) NIH Intramural Sequencing Center, 8717

Grovenmont Circle, Gaithersburg, MD 20877, USA

On Feb 6, 2003 this sequence version replaced gi:25699809.

----- Genome Center Sequencing Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.zoo@nhgri.nih.gov

----- Project Information

Center project name: crv

Center clone name: 506N12

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 119159 bases at least Q40

Consensus quality: 120104 bases at least Q30

Consensus quality: 120787 bases at least Q20

Insert size: 113000; agarose-fp

Insert size: 121109; sum-of-contigs

Quality coverage: 10.42x in Q20 bases; agarose-fp

Quality coverage: 9.72x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 6 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 18392: contig of 18392 bp in length

* 18393 18492: gap of unknown length

* 18493 25477: contig of 6985 bp in length

* 25478 25577: gap of unknown length

* 25578 46406: contig of 20829 bp in length

* 46407 46506: gap of unknown length

* 46507 49074: contig of 2568 bp in length

* 49075 49174: gap of unknown length

* 49175 89136: contig of 39962 bp in length

* 89137 89236: gap of unknown length

* 89237 121609: contig of 32373 bp in length.

FEATURES	Location/Qualifiers	
source	1. .121609	
	/organism="Bos taurus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9913"	
	/clone="RP42-506N12"	
	/clone_lib="RP42"	
misc_feature	1. .18392	
	/note="assembly_fragment"	
	clone_end:17	
	vector_side:left	
misc_feature	1. .13224	
	/note="clone overlaps with GenBank Accession Number AC137655 clone RP42-542P14 (center project name cru)"	
misc_feature	18493..25477	
	/note="assembly_fragment"	
misc_feature	25578..46406	
	/note="assembly_fragment"	
misc_feature	46507..49074	
	/note="assembly_fragment"	
misc_feature	49175..89136	
	/note="assembly_fragment"	
misc_feature	89237..121609	
	/note="assembly_fragment"	
	clone_end:SP6	
	vector_side:right	
Query Match	52.1%; Score 1862.6; DB:1; Length 121609;	
Best Local Similarity	74.8%; Pred. No. 0;	
Matches 2359;	Conservative 0; Mismatches 194; Indels 600; Gaps 8;	
QY	249 GTGGGGCCAAACCAAGATGGGAGGTCATGGTGAGAGATCTGACAGAAATGGTCCACTG	308
Db	68527 GTGGGGCCAAACCAAGACAGCGGGTCATGGTGAGAGGTCCTGACAGAAATGGTCCCTG	68586
QY	309 GAGAGGGGAAT-GCAAAACACATTCAGTATCTTGGCTTGAGAACCCCATGACAGATGA	367
Db	68587 GAGAAGGGAGTGGCAACTACTTTCAGTGTCTTGGCTTGAGAACCCCATGACAGATGA	68646
QY	368 AAAGGCAAAATGATAGGATCTGAAAGAGGAATCTCCCAAGGTGAGTGGTCCCATATG	427
Db	68647 AAAGGCAAAATGATAGGATCTGAAAGGGGAATCTCCCTGGGTGGTGGTCCCATATG	68706
QY	428 CTACTGGAGATCAGTGGAGAAATAATCTCAGAAAGAAATGAAGATGGAGCCAAAGCAAA	487
Db	68707 CTACAGGAGATGAGTGTAGAAATAATCTCAGAAAGAAATGAAGATGGAGCCAAAG-TAA	68765
QY	488 AAGATACCCAGCTGTGGATGTGACTGGTGATATAAGCAAGGTCCGATGCTGTAAGAGAC	547
Db	68766 AACAATACCAGTCTGGATGTGATGTTGATGATGAAGCAAGGTCCGATGCTGTAAGAGAC	68825
QY	548 AATATTGCATAGGAACCTTGGAAATGTGAGTCCATCAATCAAGGCAAAATGGAAATGGTCA	607
Db	68826 AATATTGCATAGGAACCTTGGAAATGTGAGTCCATCAATCAAGGCTAACTGCAAGTGTCA	68885
QY	608 AACAGAGATGCAAGATGATGTGACATCTAGGATCAGCAACTAAATGAGTCACTG	667
Db	68886 AACAGAGATGCAAGATGATGTGACATCTAGGATCAGCAACTAAATGAGTCACTG	68945
QY	668 GAATGGGTGAATTTAACTCAGATGACCATTTATCTACTCTGCGGGCAGGAATCCCTCA	727
Db	68946 GAATGGGTGAATTTAACTCAGATGACCATTTATCTACTCTGCGGGCAGGAATCCCTCA	69005
QY	728 GAAGAAATGGATAGCCATCATGCTCAACAAAGAGTCGAAATCAGTACTTGGATGCA	787
Db	69006 GAAGAAAGAGTAGCCATCATGCTCAACAAAGAGTCGAAATCAGTACTTGGATGTA	69065
QY	788 GTCTCAAAAACGACAGATGATCTCTGTTTGTTCCTCAAGGCAAAACCAATTAATATCACAG	847
Db	69066 ATCTCAAAACTGACAGATGATC-----GTTGTTTCTAAGGCAAAACCAATTAATATCATAG	69121
QY	848 TAATCCAGTCTATGCCCCCAACAGTAATGCTGAAGAGAGCTGAAGTGAACGCTCATG	907

Db 69122 TAATCCAGTCTATGCCCCAACCAAGTAATGCTGAAGAAGCTGAAGTTGAAACAGTTCTATG 69181
Qy 908 AAGACCTACAGACCTTTTGAACCTAACAACCCAAAAAGATGTCTCTCTCATTTAGGG 967
Db 69182 AAGACCTACAGACCTTTTGAACCTAACAACCCAAAAAGATGTCTCTCTCATTTAAGGG 69241
Qy 968 ACTGGAATGCAAAAGTGAAGCAAGAAACACCTGGAGTAACAGGCAAAATTTGGCCTTG 1027
Db 69242 ACTGGAATGCAAAAGTGAAGCAAGAAACACCTGGAGTAACAGGCAAAATTTGGCATTG 69301
Qy 1028 GAATACGGAATGAAGCAGGCAAGACTAATAGAGTTTTTGGCAAGAAATGCATGGTCA 1087
Db 69302 GAACACGGAATGAAGCAGGCAAGACTAATAGAGTTTTTGGCAAGAAACGCACTGGTCA 69361
Qy 1088 TAGCAAAACACCTCTTCCAAACACACAGAGAGAGCTCTACACATGGACATCACAGATG 1147
Db 69362 TAGCAAAACACCTCTTCCAAACACACAGAGAGAGCTCTACACATGGACATCACAGATG 69421
Qy 1148 GTCAACACCGAAATCAGATTGATTATATCTTTTGCAGCCAAAGATGGAGAGCTCTATAC 1207
Db 69422 GTCAACACCGAAATCAGATTGATTATATCTTTTGCAGCCAAAGATGGAGAGCTCTATAC 69481
Qy 1208 AGTCAGCAAAACACAGACAGAGCTTACTGTGGCTCAGATCATGAATCTCTTATTGCCA 1267
Db 69482 AGTTAGCAAAACACAGACAGAGCTGACTGTGGCTCAGATCATGAATCTCTTATTGCCA 69541
Qy 1268 AATTGAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGATCACTCAGGTAAGACCTAA 1327
Db 69542 AATTGAGACTTAAATTTGAAGAAAGTAGGAAACCACTAGACCGTTACGGTATGACCTAA 69601
Qy 1328 ATCCAAATCCCTTATGATTATACAGTGAAGTGAAGAAATAGATTTAAGGCGCTAGATCTGA 1387
Db 69602 ATCCAAATCCCTTATGATTATACAGTGAAGTGAAGAAATAGATTTAAGGCGCTAGATCTGA 69661
Qy 1388 TAGACAGAGTACCTTAATGAACTATGGAAGAGAGTTTATGATGTACAGAGACAGGGA 1447
Db 69662 TACATAGAGTGGCTGATGAACATATGGAATGAGTTTCGTCGATGTACAGAGACAGGGA 69721
Qy 1448 TCGAGACCATCCCAATGGAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGAGGCGCT 1507
Db 69722 TCAAGACCATCCCAATGGAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGAGGCGCT 69781
Qy 1508 TACAAATAGCTGTGAAAAGAGAGAAAGTGAAGCAAGGAAAGAAAGAAAGATATAAGC 1567
Db 69782 TACAAACATGTGTGAAAAGAGAGAAAGTGAAGCAAGGAAAGAAAGAAAGATATAAGC 69841
Qy 1568 ATCTCAATGTCAGAGTTCCAAA----- 1588
Db 69842 ATCTCAATGTCAGAGTTCCAAAAGATAGCAAGAGAAATAGAAAGCCTTCCTCAGCGACC 69901
Qy 1589 ----- 1588
Db 69902 AGTGCAAAAGAAACAGAGGAAACAAATAGAATGGGAAGACTAGAGATCTCTTCAAAAAA 69961
Qy 1589 ----- 1588
Db 69962 TTAGAGATACCAAGGGAACATTTTCATGCAAGATGGGCTCGATAGAGACAGAAATGGTA 70021
Qy 1589 ----- 1588
Db 70022 TGCACCTAAAAGAGCTGACAATATTAAGAGAGGTGGCCAGAAATACACAGAAACCGT 70081
Qy 1589 ----- 1588
Db 70082 ACAAAAAAGATCTTTTACAACCAAGATATACGATGGTGTGATCACTGACCTAGAGCCAG 70141
Qy 1589 ----- 1588
Db 70142 ACATCCTGGAACGTGAAGTCAAGTGGGCTTTAGAAAGCATCACTATGAACAAAGCTAGTG 70201
Qy 1589 ----- 1588
Db 70202 GAGGTGATGGAATTCAGTTGAGCTATTTCAAATCCTGAAAGCATGATCCTGTGAAAGTGC 70261

Qy 1589 ----- 1588
Db 70262 TGCACCTCAATATGCCAGCAAAATTTGGAAAACTCAGCAGGGGCCACAGGACTGGAAAAAGT 70321
Qy 1589 ----- 1588
Db 70322 CAGTTTTTCATTTCAATCCCAAGAAAGCAATGCCAAAGAAATGCTCAAACTCCACATAA 70381
Qy 1589 ----- 1588
Db 70382 TTGCACCTCATCTCACACACTAGGAAGTAATGCTCAAAATTTCTTCAAGCCAGGTTTCAGC 70441
Qy 1589 ----- 1641
Db 70442 AATATGTGAACCTACAGATGTTTAAAGCTGGTTTTAGAAAAAGGAGAGGAAACAGATCA 70501
Qy 1642 AATTGCCAACATCTCTGTATATGAAAGAAAGCAAGAGAGTTCCAGAAAAACATCTATTT 1701
Db 70502 AATTGCCAACATCCACTGGATTTACAAAAGAGAGAGAGTTCCAGAAAAACATCTATTT 70561
Qy 1702 CTGCTTTATTTGACTATGCAAAAAGCCTTTGACTGTGGGCTCACAATAAACTGTGAAAAAT 1761
Db 70562 CTGCTTTATTTGACTATGCAAAAAGCCTTTGACTGTGGATCAACAATAAACTATGGAAT 70621
Qy 1762 TCTGAAAGGATGGGAATACCAAGCACCTGACCTGACTCTTGAAAAATTTGTATCGAG 1821
Db 70622 TCTGAAAGATGGGAATACCAAGCACCTGACTGCTGCTCTTGAGAAACCTGTATGAGG 70681
Qy 1822 TCAGGAAGCAACAGTTAGAACTGGACATGGAAACAACAGACTGGTTCCAAAGTAGAAAAAG 1881
Db 70682 TCAGGAAGCAACAGTTAGAACTGGACATGGAAACAACAGACTGGTTCCAAATAGAAAAAG 70741
Qy 1882 AGTATGTCAGGCTGTATATTTGTCACCGGCTTGTGTTAACTTCTATGCGAGAG-ACATCAT 1940
Db 70742 AGTATGTCAGGCTGTATATTTGTCACCGCTGCTTATTTAACTTATATGCGAGATACAT 70801
Qy 1941 GAGAAACGCTGGGCTGGAAAGCAACAAGCTGGAATCAAGATTGCCGGAGAAATAGCAA 2000
Db 70802 GAGAAACGCTGGAGAGAGAAACAACAAGCTGGAATCAAGATTGCCGGAGAAATCTCAA 70861
Qy 2001 TAACTCTCAGATATGCGAGATATACCACTCTTATGGCAGAAAGTGAAGAGAACTAAAAAG 2060
Db 70862 TAACTCTCAGATATGCGATATACCACTCTTATGGCAGAAAGTGAAGAGAACTAAAAAG 70921
Qy 2061 CCTCTTGTGAAGGTGAAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAA 2120
Db 70922 CCTCTTGTGAAGGTGAAAGAGAGAGAAAGTTGGCTTAAAGCTCAACATTCAGAA 70981
Qy 2121 AACGAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAATAGATGGGGAACACGTGGA 2180
Db 70982 AATGAAGATCATGGTATCTGGTCCCATCACTTCATGGGAAAGAGATGGGGAACACGTGGA 71041
Qy 2181 AACAGTGTGAGCTTTATTTTGGGGGCTCCAAATCACTGCAGATGGTGCAGCTCAGC 2240
Db 71042 AACAGTGTGAGCTGTATTTTTT-ITGCTCTCCAAATCACTGCAGATGGTGTGCAAC 71100
Qy 2241 CATGAAATTAAGAGACACTTACTCTTGGAAAGAAAGTTA--ACCAACCTAGATAGCAT 2298
Db 71101 CATGAAATTAAGAGAGCTTACTCTTGGAGGAAGTTATGACCACTAGATAGCAT 71160
Qy 2299 TTGAAAGCAGAGACATTAACCTTGCACAACAGCCCATCTCTAGTCAAGGCTATGGTTTT 2358
Db 71161 TTCAAAAGCAGAGACATTAACCTTGCACAACAAAGTTCTTCTAGTCAAGGCTGTGTTTT 71220
Qy 2359 CCAGTGGTCACTGTATGTGATGTGAGAGTTGAGCTGTGAAGAAAGCTGAGCACTGAAGAT 2418
Db 71221 CCAGTGTATGTATGTGATGTGAGAGTTGAGCTGTGAAGAAAGCTGAGCGCTGAAGAT 71280
Qy 2419 GATGCTTTTGAAGTGTGGTGTGGAGAGACTCTTTGAGAGTCCCTTTGGACTGCAAGGAGA 2478
Db 71281 GATGCTTTTGAAGTGTGGTGTGGAGAGACTCTTTGCGAGTCCCTTTGGACTGCAAGGAGA 71340

2479	TCCAAACGATCCATTCTCAAGAGATCAGCCCTGGGATTTCTTTGGAGGAATGATGCTA	2538
71341	TCCAAACGATCCATCCTAAAGAGATCAGTCTGCTGTTCTTTGGAGGACTTTTCTGT	71400
2539	ANGCTGAACTCCCACTACTTTTGGCCACCTGATCAGAGAGCTGACTCAGTGGAAAAGACC	2598
71401	ANGCTGAACTCCCAATACTTTTGGCCACCTCATGCGAGAGTTGACTCATTTGGAAAAGACT	71460
2599	CTGATGCTGGGAGGATTTGGGGCAGGAGGACAGGGGACACAGAGATGAGATGGCTG	2658
71461	CTGATGCTGGGAGGATTTGGGGCAGGAGGACAGGGGACACAGAGATGAGATGGCTG	71520
2659	GATGGCATCAGTACTCCATGAGCATGAGTCTGGGTGAATCTCTGGAGTTGGTGATGGAC	2718
71521	GATGGCTTCCGACACACAATGGACATGATTTCACTGAACCTCGGGAGTTGGTGATG--	71577
2719	AGGAGGCTGTCTCGCGGATTCATGGGTACAAAGAGTTGGACAGCTCAGCAAC	2778
71578	-GTGAGGCTGGGTTACTGCAACTCATGGGTACAAAGAGTCCAAACAGCTGAGTGAC	71636
2779	TGAAGTGAAGTGAAGTGTACTGAAACCTTAGTA	2811
71637	TGAACTGAAGTGAAGTGTACTGAAACCTTAGTA	71669

RESULT 45	AC140093	155001 bp	DNA	linear	HTG 22-MAR-2003
AC140093/C	Bos taurus	clone RP42-362C3,	WORKING DRAFT	SEQUENCE, 8 ordered	
LOCUS	pieces				
DEFINITION	AC140093				
ACCESSION	AC140093.2	GI:29150354			
VERSION	HTG: HTGS PHASE2: HTGS DRAFT,				
KEYWORDS					

Bos taurus (Cow)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 155001)
 REFERENCE
 AUTHORS
 Akhter N., Antonellis A., Ayele K., Beckstrom-Sternberg S.M.,
 Ben-Jamin B., Blakesley R.W., Bouffard G.G., Brinkley C., Brooks S.,
 Cariaga K., Coleman B., Engle J., Granite S., Guan X., Gupta J.,
 Haghghi P., Han J., Hansen N., Ho S.-L., Idol J.R., Karlins E.,
 Laric P., Lee-Lih S.-Q., Legaspi R., Maduro Q.L., Maduro V.B.,
 Marques S.H., Mastello C., Maskeri B., McDowell J.,
 Parguligan C., Pearson S., Portnoy M.E., Prasad A.,
 Reddi-Dugue N., Schandler K., Schueler M.G., Sison C.,
 Stantripop S., Thomas J.W., Thomas P.J., Touchman J.W., Vogt J.L.,
 Wetherby K.D., Wiggins L., Young A. and Green E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 155001)
 REFERENCE
 AUTHORS
 Green E.D.
 TITLE
 Submission
 Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 155001)
 REFERENCE

Direct Submission
 Submitted (22-MAR-2003) NIH Intramural Sequencing Center, 8717
 JOURNAL Government Circle, Gaithersburg, MD 20877, USA
 COMMENT On Mar 22, 2003 this sequence version replaced gi:28460769.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc_zoo@nhgri.nih.gov
 ----- Project Information
 Center project name: dds
 Center clone name: 262C03

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated

order and orientation of each sequence contig has been reestablished using one or more of the following: read-pair data from individual subclones, overlaps with neighboring subclones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- summary statistics -----

Sequencing vector: plasmid, n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 153316 bases at least Q40
Consensus quality: 153884 bases at least Q30
Consensus quality: 154212 bases at least Q20
Insert size: 138000; agarose-fp
Insert size: 154301; sum-of-contrigs
Quality coverage: 9.57% in Q20 bases; agarose-fp
Quality coverage: 8.56% in Q20 bases; sum-of-contrigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

*	1	719:	contig of 719 bp in length
*	720	819:	gap of unknown length
*	820	14217:	contig of 13398 bp in length
*	142318	14317:	gap of unknown length
*	14318	14197:	contig of 1880 bp in length
*	16198	12897:	gap of unknown length
*	16298	19432:	contig of 2135 bp in length
*	18433	19532:	gap of unknown length
*	18533	32260:	contig of 13728 bp in length
*	32261	32360:	gap of unknown length
*	32361	78744:	contig of 44384 bp in length
*	76745	108645:	gap of unknown length
*	76845	109615:	contig of 32771 bp in length
*	109616	109715:	gap of unknown length
*	109716	155001:	contig of 45286 bp in length.

FEATURES	Location/Qualifiers
105/10	150001: 1000000

```

1. .155001
source      /organism="Bos taurus"
            /mol_type="genomic DNA"
            /db_xref="taxon:9913"
            /clone="RP42-262C3"
            /clone_lib="RP42"
            1. 75001
misc_feature /notes="clone overlaps with GenBank Accession Number
AC139631 clone RP42-108E2 (center project name ddr)"
            1. .719
misc_feature /notes="assembly_fragment
            clone end:T7
            vector side:left"
            820..14217
misc_feature /notes="assembly_fragment"
            14318..16197
misc_feature /notes="assembly_fragment"
            16298..18432
misc_feature /notes="assembly_fragment"
            18533..32260
misc_feature /notes="assembly_fragment"
            32361..76744
misc_feature /notes="assembly_fragment"
            76845..109615
misc_feature /notes="assembly_fragment"
            109716..155001

```


/note="assembly_fragment
clone end:396
vector_side:right"

Query Match	52.1%	Score 1861.5;	DB 1;	Length 155001;
Best Local Similarity	75.0%;	Pred. No. 0;		
Matches 2331;	Conservative 0;	Mismatches 215;	Indels 561;	Gaps 6;
QY 358	AACAGTATGAAAGCAAAATGATAGGATAGTAAAGAGGAACCTCCCGAGTCTAGTAGGT	417		
Db 122336	AAAAGTATGAAAGGCAAAATGATAGGATAGTAAAGAGGAACCTCCCGAGTCTAGTAGGT	122277		
QY 418	GCCCATATGCTACTGGAGATCAGTGGAGAAATTAATCTCCAGAAAGAAATGAAGAGATGGAG	477		
Db 122276	GCCCAATGTCTACTGGAGATCAGTGGAGAAATTAATCTCCAGAAAGAAATGAAGAGATGGAG	122217		
QY 478	CCAAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGGTGTATATAAGCAAGGTCCGATGC	537		
Db 122216	CCAAAGCAAAAGAAATACCCAGCTGTGGTGTGACTGGTGTATATAAGCAAGGTCCGATGC	122157		
QY 538	TGTAAGAGCAATATTTGATAGGAACCTGGAAATGTCAGGTCCATGAATCAAGGCAAAATTG	597		
Db 122156	TGTAAGAGCAATATTTGATAGGAACCTGGAAATGTCAGGTCCATGAATCAAGGCAAAATTG	122097		
QY 598	GAAATGGTCAAAACAGAGATGCAAGAGTGAATGTCAACATTCTAGGAATCAGCGAACTA	657		
Db 122096	GAAATGGTCAAAACAGAGATGCAAGAGTGAATGTCAACATTCTAGGAATCAGCGAACTG	122037		
QY 658	AAATGGATGGAAATTTAACTCAGATGACCAATTAATCTACTCTGCGGCGAG	717		
Db 122036	AAATGGATGGAAATTTAACTCAGATGACCAATTAATCTACTCTGCGGCGAG	121977		
QY 718	GAAATCCCTCAGAAGAAATGGAGTACCATCATGTCAACAAAGAGTCCGAATGCCAGTA	777		
Db 121976	GAAATCCCTCAGAAGAAATGGAGTACCATCATGTCAACAAAGAGTCCGAATGCCAGTA	121917		
QY 778	CTTGATCGAGTCTCAAAAACGACAGAGATGATCTCTGTCTTTTCCAAAGCAAAACCATTC	837		
Db 121916	CTTGATCGAGTCTCAAAAACGACAGAGATGATCTCTGTCTTTTCCAAAGCAAAACCATTC	121857		
QY 838	AATATCAGATTAATCCAGTCTATGCCCAACCAAGTAAATGCTGAAGAGCTGAAGTTGAA	897		
Db 121856	AATATCAGATTAATCCAGTCTATGCCCAACCAAGTAAATGCTGAAGAGCTGAAGTTGAA	121797		
QY 898	CGGTCTATGAAGACCTCAAGACCTTTTGAACCTAACACCCCAAAAAGAGTGTCTTCTC	957		
Db 121796	CAGTCTCTGAAGACCTCAAGACCTTTTGAACCTAACACCCCAAAAAGAGTGTCTTCTC	121737		
QY 958	ATTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAAGAAACACCTGGAGTAAACAGGCAAA	1017		
Db 121736	ATTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAAGAAACACCTGGAGTAAACAGGCAAA	121677		
QY 1018	TTTGGCCCTTGGAAATACGGAATGAAGCGGCAAGAGCTAATAGAGTTTTCGCAAGAAAT	1077		
Db 121676	TTTGGCCCTTGGAAATGAAGAGCGGCAAGAGCTAATAGAGTTTTCGCAAGAAAT	121617		
QY 1078	GCATGTGCTATAGCAAAACCCCTCTTCCAAACCAACAGAGAGAGACTCTACATGGACA	1137		
Db 121616	GCATGTGCTATAGCAAAACCCCTCTTCCAAACCAACAGAGAGAGACTCTACATGGACA	121557		
QY 1138	TCACAGATGTCAACACCGAAATCAGATTGATTAATCTTTTCAGCGCAAAAGATGGAGA	1197		
Db 121556	TCACAGATGTCAACACCGAAATCAGATTGATTAATCTTTTCAGCGCAAAAGATGGAGA	121497		
QY 1198	AGCTCTATACAGTCAAGCAAAACCAAGACCGAGGCTTACTGTGCTCAGATCATGAATC	1257		
Db 121496	AGCTCTATACAGTCAAGCAAAACCAAGACCGAGGCTTACTGTGCTCAGATCATGAATC	121437		
QY 1258	CTTATTGCCAAATTCAGACTTAATTTGAAGAAAGTAGGAAAAACCACTAGATCACTCAGG	1317		
Db 121436	CTTATTGCCAAATTCAGACTTAATTTGAAGAAAGTAGGAAAAACCACTAGATCACTCAGG	121377		
QY 1318	TAAGACCTAAATCCAAATCCCTTATGATTAATACAGTGGAAAGTGAAGATAGATTAAAGGC	1377		

Db 121376	TATGACCTAAACTAAATCACTTAAGATTATACAGTGGAACTGAGAAATAGATTAAAGGC	121317		
QY 1378	CTAGATCTGATAGACAGTACCTTAATGACTTATGACAGAGGTTTATGACATGTACAG	1437		
Db 121316	CTAGATCTGATAGATAGTGCCTGATGAATATGAATGAGGTTCTGTGATATTGTACAG	121257		
QY 1438	GAGACAGGGATCGAGACCATCCCTCATGGAAAAAGAAATGCAAAAAAGCAAAATGGCTGTCT	1497		
Db 121256	GAGACAGGGATCAAGACCATCCCTCATGGAAAAAGAAATGCAAAAAAGCAAAATGGCTGTCT	121197		
QY 1498	GGGGAGCCCTTACAAATAGCTGTGAAAAAGAGAAAGTGAAGAAAGCAAGAAAAAGGAA	1557		
Db 121196	GGGGAGCCCTTACAAATAGCTGTGAAAAAGAGAAAGTGAAGAAAGCAAGAAAAAGGAA	121137		
QY 1558	AGATAAAAGCATCTGAATGAGAGTTCCTCAAGAAATTCCTCAGTTG-----	1601		
Db 121136	AGATAAAAGCATCTGAATGAGAGTTCCTCAAGAAATTCCTCAGTTG-----	121077		
QY 1602	-----TTCAAGCTGGT 1612			
Db 121076	TTCAAGCTGGT-----TTCAAGCTGGT 1612			
QY 1613	TTTAGAAAAAGTCAAGAGAACCCAGAG-----	1637		
Db 121016	TTTAGAAAAAGTCAAGAGAACCCAGAG-----	120957		
QY 1638	-----	1637		
Db 120956	AGAAATGGTATGACCTAACAGAAAGCAGAAAGATATTAAAGAGAGTGGCAAGATATACA	120897		
QY 1638	-----	1637		
Db 120896	GAAGAACTATACAAAAAGAGATTTTCAGCAAGATAATCAAGATGTGTGATCACTGAC	120837		
QY 1638	-----	1637		
Db 120836	CTAGACGACAGACATCTGGAAATGTGAAGTCAAGTGGCCCTTAGAAAGCATCACTACGAAC	120777		
QY 1638	-----	1637		
Db 120776	AAAGCTAGGGAGGTGATGGAATTCAGTGGAACTATTTCAAATCCTGAAAGATGATGCT	120717		
QY 1638	-----	1637		
Db 120716	GTGAAAGTGTGACCTCAATATGCAAGCAAAATTTGGAAAACTCAGCAGTGGCCACAGGAC	120657		
QY 1638	-----	1637		
Db 120656	TGAAAAAGGTCAATTTTCATTTCCAAATCCCAAGAAATGCTCAAACTACCGCACAATTGCAC	120597		
QY 1638	-----	1637		
Db 120596	TCATCTCACACGCTAGTAAAGTAATGCTCAAAATTTCTCAAGTCACTTTCAGCAATACG	120537		
QY 1638	-----ACCAATTTGCCAAACATCTCTGTATC 1663			
Db 120536	TGAACCATGAATCTCCAGATGTTTCAAGCTGGTTTAGAAAAAGCCCAATCCGCTGGATC	120477		
QY 1664	ATGAAAAAGCAAGAGAGTTCAGAAAAACATCTATTTCTGCTTTTATTGACTATGCAAAA	1723		
Db 120476	ATGAAAAAGCAAGAGAGTTCAGAAAAACATCTATTTCTGCTTTTATTGACTATGCAAAA	120417		
QY 1724	GCCTTTGACTGTGGGGTCACAATTAACATGCTGAAAAATTTCTGAAAGGGATGGAATACCA	1783		
Db 120416	GCCTTTGACTGTGGGGTCACAATTAACATGCTGAAAAATTTCTGAAAGGGATGGAATACCA	120357		
QY 1784	GACCACTGACCTGACTCTTTGAAAAATTTGATGAGCTCAGGAAAGCAACAGTTAGAATC	1843		
Db 120356	GACCACTGACCTGACTCTTTGAAAAATTTGATGAGCTCAGGAAAGCAACAGTTAGAATC	120297		
QY 1844	GGACATGGAAACACAGACTGTTTCCAAAGTAGGAAAGGATGATGTCAGGCTGTATATG	1903		

Db 120296 GGACATGCAACACAGACTGGTTCCAAATAGGAAAGAGTACGTCAAGGCTGTATATTG 120237
Qy 1904 TCACCCGGCTTGTTTAACTTCTTATGACAG-ACATCATGAGAAACGCTGGGCTGGAAGAA 1962
Db 120236 TCACCCCTGCTTATTTAATCTTATATGACAGTACATCATGAGAAACACTGGGCTGGAAGAA 120177
Qy 1963 GCACAAGCTGGAAATCAAGATTCGCGGAGAAATAGCAATACCTCAGATATGCAGATGAT 2022
Db 120176 GCACAAGCTGGAAATCAAGATTCGCGGAGAAATATCAATAACCTCAGATATGCAGATGAC 120117
Qy 2023 ACCACCTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCTCTTGATGAAGGTCAAGAG 2082
Db 120116 ACCACCTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCTCTTGATGAAGGTCAAGAG 120057
Qy 2083 GAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAAGAGATCATGGCATCTGGT 2142
Db 120056 GAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAAGAGATCATGGCATCTGGT 119997
Qy 2143 CCCATCACTTCATGGGAATATAGATGGGAAACAGTGGAAACAGTGTCAAGCTTTATTTT 2202
Db 119996 CCCATCACTTCATGGGAATATAGATGGGAAACAGTGGAAACAGTGTCAAGCTTTATTTT 119937
Qy 2203 TGGGGGGTCCAAAATCACTGCAGATGTGACTGCAGCCATGAAATTAAGAGACACTTAC 2262
Db 119936 --GGGGGTCCAAAATCACTGCAGATGTGACTGCAGCCATGAAATTAAGAGAGCTTAC 119879
Qy 2263 TCCTTGGAGAAAGTTA--ACCACCTAGATAGCATATTCGAAAGCAGACATTAACCT 2320
Db 119878 TCCTTGGAGAAAGTTATGACCAACCCAGATAGCATATTCGAAAGCAGACATTAACCT 119819
Qy 2321 TGCCAAACAAAGCCCATCTAGTCAAGGCTATGGTTTTCCAGTGGTCAATGATGATGTG 2380
Db 119818 TGCCAAACAAAGGTTCACTAGTCAAGGCTATGGTTTTCCAGTGGTGTGATGGATGTG 119759
Qy 2381 AGAGTTGAGCTGTGAAGAAAGTGAAGCTGAAGTAATGATGCTTTGAACTGTGGTGT 2440
Db 119758 AGAGTTGGGCTGTGAAGAAAGTGAAGCTGAAGTAATGATGCTTTGAACTGTGGTGT 119699
Qy 2441 GCAGAAAGCTCTTGAAGTCCCTTGGACTGCAAGAGATCCAAACAGTCCATCTGAAGG 2500
Db 119698 G--GAAACTCTTGAAGTCCCTTGGACTGCAAGAGATCCAAACAGTCCATCTGAAGG 119641
Qy 2501 AGATCAGCCCTGGGATTTCTTGGAAAGAAATGATCTAAAGCTGAACCTCAGTACTTTG 2560
Db 119640 AGATCAGCCCTGGGATTTCTTGGAAAGGTTGATCTGAAGCTGAACCTCAGTACTTTG 119581
Qy 2561 GCCACTGATCAGAAAGCTGACTCACTGGAAGACCCCTGATGCTGGAGGGATTTGGG 2620
Db 119580 GCCACTCATGAAGAGAGTTAACTCATTTGAAAGACTCTGATGCTGGAGGGATTAGGG 119521
Qy 2621 GCAGGAGAGAGGGGACGACAGAGATGAGATGGCTGGATGGCATCACTGATGATGG 2680
Db 119520 GCAGGAGAGAGGGGACGACAGAGATGAGATGGCTGGATGGCATCACTGATGATGG 119461
Qy 2681 AGTGAAGTCTGGTGAATCTCCTGGAGTGGTATGATGACAGAGGAGGCTCTCTGCGGGA 2740
Db 119460 ACATGAGTTTGAAGTGAATTCGCGGAGTGGTATGACAGGAGGAGGCTGGTGTCTGCGA 119401
Qy 2741 TTCATGGGCTCAAGAGTGGACACCACTGAGCAACTGAGCAACTGAACCTGAACCTGACTG 2800
Db 119400 TTCATGGAGTCAACAGAGTCCGACACGACTGAGCGAGTGAACCTGAACCTGAACCTG 119341
Qy 2801 AAACCTTAGTATTATATCTAGAAATAGTAAATTTTATATGATGATGATGATGATGATG 2860
Db 119340 TGTACTGTTGACAGGAGGAGTAAAGCTGGTATAGGCATTATGGAAGAGCTGAGGATTTT 119281
Qy -2861 CATAATGTTGGTTAAGATAATAGATTTTCAAAATGATTTTATCTT 2907
Db 119280 TCAAAAAATACAAATAGAACTACCATATATACAGGAATTTCACTT 119234

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 13:00:13 ; Search time 650 Seconds
(without alignments)
3.920 Million cell updates/sec

Title: US-10-664-775-2

Perfect score: 3572

Sequence: 1 gtcaggagggcgcagtga.....gcaacaacagcagaagctt 3572

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 0.5

Searched: 41 seqs, 356661 residues

Total number of hits satisfying chosen parameters: 82

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rng2.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
C 1	1842.4	51.6	54842	1	ABS55699
C 2	1051.2	29.4	38793	1	ABS53700
C 3	1028.7	28.8	17141	1	AAD26486
C 4	1012.1	28.3	78056	1	ABV93701
C 5	1009	28.2	10492	1	AAZ58843
C 6	966.804	27.1	23653	1	ABK14793
C 7	918.101	25.7	11288	1	ABV72459
C 8	917.7	25.7	10492	1	AAZ58843
C 9	885.401	24.8	78056	1	ABV93701
C 10	672	18.8	3370	1	AAQ79534
C 11	667.9	18.7	2830	1	AAT96051
C 12	619.909	17.4	54842	1	ABS55699
C 13	589.4	16.5	6620	1	ABV72460
C 14	557.9	15.6	23653	1	ABK14793
C 15	494.6	13.8	13723	1	AAV74739
C 16	413.4	11.6	908	1	ABN74558
C 17	406.2	11.4	1333	1	AAH21005
C 18	388.9	10.9	1813	1	AAD26492
C 19	380.6	10.7	967	1	ADE81111
C 20	372.6	10.4	419	1	ABX48667
C 21	364.7	10.2	451	1	ABX35679
C 22	363.2	10.2	436	1	ABX40926
C 23	360.3	10.1	432	1	ABX39926
C 24	357.1	10.0	433	1	ABX35919
C 25	353.3	9.9	433	1	ABX40182
C 26	348.7	9.8	1848	1	AAH21025
C 27	347.4	9.7	436	1	ABX41993
C 28	344.199	9.6	31412	1	ABQ75995
C 29	344.199	9.6	31412	1	ABV93702
C 30	342.4	9.6	11034	1	AAD26491
C 31	335.8	9.4	428	1	ABX42989
C 32	333	9.3	406	1	ABX36143
C 33	326.3	9.1	1529	1	AAH21026

C 34	324	9.1	423	1	ABX43146	Bovine EST associa
C 35	321.6	9.0	410	1	ABX48669	Bovine EST associa
C 36	320.8	9.0	423	1	ABX40334	Bovine EST associa
C 37	320.4	9.0	409	1	ABX49143	Bovine EST associa
C 38	316.4	8.9	529	1	AAZ56452	Human immune/baema
C 39	313.4	8.8	425	1	ABX44380	Bovine EST associa
C 40	312.4	8.7	1435	1	AAH21024	Bovine-derived DNA
C 41	307.8	8.6	422	1	ABX38226	Bovine EST associa
C 42	306.2	8.6	3649	1	AAH50416	Calf acetyl cholin
C 43	299.8	8.4	336	1	ABX46049	Bovine EST associa
C 44	299.4	8.4	447	1	ABX46780	Bovine EST associa
C 45	298.6	8.4	815	1	AAD26490	Sheep alpha (1, 3)

ALIGNMENTS

RESULT 1

ABS55699/C

ID ABS55699 standard; DNA; 54842 BP.

XX ABS55699;

AC ABS55699;

XX 08-JAN-2003 (first entry)

DT 08-JAN-2003 (first entry)

XX Bovine Claudin-16 deficiency associated polynucleotide sequence #2.

DE Bovine Claudin-16 deficiency associated polynucleotide sequence #2.

XX Bovine; Claudin-16 deficiency; type 2 mutation; ds.

KW Bovine; Claudin-16 deficiency; type 2 mutation; ds.

XX Bos sp.

OS Bos sp.

PN JP2002238570-A.

XX 27-AUG-2002.

PD 27-AUG-2002.

XX 14-FEB-2001; 2001JP-00037623.

PF 14-FEB-2001; 2001JP-00037623.

XX 14-FEB-2001; 2001JP-00037623.

PR 14-FEB-2001; 2001JP-00037623.

XX (CHIK-) CHIKUSAN GIJUTSU KYOKAI SH.

PA (KACH-) KACHIKU KAIRYO JIGYODAN SH.

PA (DOKU-) DOKURITSU GYOSEI HOJIN KACHIKU KAIRYO CE.

XX WPI; 2002-744759/81.

XX Gene diagnosis of type 2 mutation for bovine claudin-16 deficiency, comprises checking for mutation in 1-161 position in claudin-16 sequence.

PT Example 1; Page 9-28; 43pp; Japanese.

XX The invention describes gene diagnosis of type 2 mutation for bovine claudin-16 deficiency. The method involves (a) getting a bovine nucleic acid sample; (b) subjecting the sample to a gene amplification reaction to give a nucleic acid fragment; (c) checking the presence of mutation in the nucleic acid fragment of step (b). The method is used for gene diagnosis of type 2 mutation for bovine claudin-16 deficiency. This sequence represents a polynucleotide from the bovine claudin-16 deficiency associated gene in which the type 2 mutation occurs

XX Sequence 54842 BP; 15478 A; 10714 C; 10200 G; 18336 T; 0 U; 114 Other;

XX Query Match 51.6%; Score 1842.4; DB 1; Length 54842;

XX Best Local Similarity 75.6%; Pred. No. 4.8e-05;

XX Matches 2295; Conservative 0; Mismatches 131; Indels 608; Gaps 7;

XX 358 AACAGTATGAAAGCAAAATCATAGGATCTGAAAGAGGAACTCCACAGGTCAGTAGGT 417

XX 16137 AACATATGAAAGCAAAATCATAGGATCTGAAAGAGGAACTCCACAGGTCAGTAGGT 16078

XX 418 GCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGATGAAGATGGAG 477

XX 16077 GCCCAATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGGAGAGTAAAGGATGGAG 16018

QY 478 CCAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGGTGATATAAGCAAGGTCGGATGC 537
Db 16017 CCAAGCAAAAGAAATACACAGCTGGGATGTGACTGGTGATAGAAGCAAGGTCGGATGC 15958
QY 538 TGTTAAGGCAATATTGCGATAGGAACCTGGATGTGAGTCCATGATCAAGCAAAATTG 597
Db 15957 TGTTAAGGCAATATTGCGATAGGAACCTGGATGTGAGTCCATGATCAAGCAAAATTG 15898
QY 598 GAAGTGGTCAAAAGAGATGCCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAACTA 657
Db 15897 GAAGTGGTCAAAAGAGATGCCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAACTG 15838
QY 658 AAATGGACTGGAATGGGTGAATTTAACTCAGATGACCAATTTATATCTACTACTGCGGCAG 717
Db 15837 AAATGGACTGGAATGGGTGAATTTAACTCAGATGACCAATTTATATCTACTACTGCGGCAG 15778
QY 718 GAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAGAGTCCGAATTCAGTA 777
Db 15777 GAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAGAGTCTGAAATGCAGTA 15718
QY 778 CTTGGATGCACTCTCAAAAACGACAGATGATCTCTGTTGTTTCCAAAGGCAAAACCAATTC 837
Db 15717 CTTGGATGCAATCTCAAAAATGACATATATGATCTCTGTTGTTTCAAGGCMAAACCAATTC 15658
QY 838 AATATCACAGTAATCCAGTCTATGCCCAACCAAGTAATGCTGAAGAGCTGAAGTTGAA 897
Db 15657 AATATCACAGTAATCCAGTCTATGCCCAACCAAGTAATGCTGAAGAGCTGAAGTTGAA 15598
QY 898 CGTCTATGAAGACCTCAAGACCTTTTAGAATTAACACCCCAAAAAGAGTCTCTTCTC 957
Db 15597 CAGTCTATGAAGACCTCAAGACCTTTTAGAATTAACACCCCAAAAAGAGTCTCTTCTC 15538
QY 958 ATTATAGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGAGTAACAGGCAAA 1017
Db 15537 ATTATAGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGAGTAACAGGCAAA 15478
QY 1018 TTTGGCCCTTGAATACGGAATGAGCAGGCAAGAGCTAATAGAGTTTGGCCAGAAAT 1077
Db 15477 TTTGGCCCTTGAATACAGANTGAAGCAGGCAAGAGCTAATAGAGTTTGGCCAGAAAT 15418
QY 1078 GCATGGTCTATAGCAAAACCCCTCTTCCAAACAACAAGAGAACTCTACACATGGACA 1137
Db 15417 GCATGGTCTATAGCAAAACCCCTCTTCCAAACAACAAGAGAACTCTATACATGGACA 15358
QY 1138 TCACAGATGTCACACCGGAATCAGATGATGATTTATTTTCGACCAAGAGTGGAGA 1197
Db 15357 TCACAGATGTCACACCTGCAACTGAAATCAGATGATGATTTATTTTCGACCAAGAGTGGAGA 15298
QY 1198 AGCTCTATACAGTACGCAAAACCAAGACCAAGAGCTTACTGTGGCTCAGATCATGAATC 1257
Db 15297 AGCTCTATACAGTACGCAAAACCAAGACCAAGAGCTGACTGTGGCTCAGATCATGAATC 15238
QY 1258 CTTATTGCAAAATCAGACTTTAAATTTGAAGAGTGAAGGAAACCACTAGATCACTCAGG 1317
Db 15237 TTTATTGCAAAATCAGACTTTAAATTTGAAGAGTGAAGGAAACCACTTAAACCACTCAGG 15178
QY 1318 TAAGACCTTAAATCCAAATCCCTTATGATTTATACAGTGAAGTGAAGAAATAGATTTTAAGGC 1377
Db 15177 TATGACCTTAAATCAATCCCTTATGATTTATACAGTGAAGTGAAGAAATAGATTTTAAGGC 15118
QY 1378 CTAGATCTGATAGACAGAGTACTTAATGAATATGAGCAGAGGTTTCATGATTTGTACAG 1437
Db 15117 CTAGATCTGATAGACAGAGTACTTAATGAATATGAGTGAAGTGAAGTTCGTGACATTTGTGAC 15058
QY 1438 GAGACAGGATCGACACCATCCCATGAAAGAAATGCAAAAAGCAAAATGCTGTCT 1497
Db 15057 GAGACAGATTCGACACCATCCCATGAAAGAAATGCAAAAAGCAAAATGCTGTCT 14999
QY 1498 GGGAGGCTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAGGAAAGGAA 1557
Db 14998 GGGAGGCTTACAAATAGCTGTGAAAAGAGAGAGTGAAGCAAGGAAAGGAA 14939
QY 1558 AGATAAAGCATCTGAATGCAGAGTTCCAAA----- 1588

Db 14938 AGATCAAGCATCTGAATSCAGAGTTCCAAAGAAATAGCAAGAGATAAGAAAGCCTTC 14879
QY 1589 ----- 1588
Db 14878 CTCAGTGATCAATGCAAGAAATAGAGGAAWCAAGATGGGAAAGACTAGANATCTC 14819
QY 1589 ----- 1588
Db 14818 TTCAAGAAAAATTAGAGATACCAAGGGAATATTTCATGCAAGATGGGCTCGATAAGGAC 14759
QY 1589 ----- 1588
Db 14758 AGAAATGGTATGATCTTAACAGACGACAGAGATATTAAAGAGAGTGGCAAGANTACACA 14699
QY 1589 ----- 1588
Db 14698 GAAGAACTGGACAAACAAAGATCTTCATGACCCAGATAATCACGATGGTGTGATCAGTGAC 14639
QY 1589 ----- 1588
Db 14638 CCAGAGCCAGAAATCCTGGAATGTGAAGTCAAGTGGGCTTAGAAAGCATCATTATGAAC 14579
QY 1589 ----- 1588
Db 14578 AAAGCTAGTGAAGTGAATCCAGTTCAGCTATTTCAGTCTCTGAAGGATGATGCT 14519
QY 1589 ----- 1588
Db 14518 GTGAAAGTCTGCATCAATATGCAGCAAAATTTGGAAAACTCAGCAGTGGCCACAGGAC 14459
QY 1589 ----- 1588
Db 14458 TGGGAAAGGTCACTTTTCATTTCCATCCCAAGAAAGCAATGCTCAAGAAATGCTCAAA 14399
QY 1589 ----- 1588
Db 14398 CTACCACACATGCATCTCATCCACACTAGTAAAGTAATGCTCAAAATTTCTCAAGCC 14339
QY 1589 -----GAACTTCAGTTGTTCAAGCTGGTTTGTAGAAAAGTC 1624
Db 14338 AGCTTCAGCAATACGTGAACCGTGAATTTCTGATGTTCAAGCTGGTTTGTAGAAAAGC 14279
QY 1625 AGAGGAACCAAGAGCAAAATTCGCAACATCTCTGATATGGAAGAAAGCAAGAGATTC 1684
Db 14278 AGAGGAACCAAGAGTCAAAATTCGCAACATCTCTGATGATGGAAGAAAGCAAGAGATTC 14219
QY 1685 CAGAAAAATCTATTTCTGCTTTATGACTATGCAAAAGCCTTTGACTGTGGGGTCAAC 1744
Db 14218 CAGAAAAATCTATTTCTGCTTTATGACTATGCAAAAGCCTTTGACTGTGGATCAC 14159
QY 1745 AATAAACTGTGAAAAATTTCTGAAAGGGATGGAATACCAAGCACCTGACCTGACTCTTG 1804
Db 14158 AATAAACTGTGAAAAATTTCTGAAAGGGATGGAATACCAAGCACCTGATCTGCTCTTG 14099
QY 1805 AAAATTTGTATGCAAGTCAGGAACCAAGTTAGAACTGGACATGGAACACAGCTGG 1864
Db 14098 AGAAATTTGTATGCAAGTCAGGAACCAAGTTAGAACTGGACATGGAACCAACAGCTGG 14039
QY 1865 TTCCAAGTAGGAAAAAGGAGTATGTCAAGGCTGTATATTGTCAACCGGCTTTGTTAACTTC 1924
Db 14038 TTCCAATAGGAAAAAGGAATACATCAAGCTGTGATTGTCAACCTGCTTACATAACTTA 13979
QY 1925 TATGCAGAG-ACATCATGAGAAACGCTGGCTGGAGGAAGCAAGCAAGCTGGAATCAAGATT 1983
Db 13978 TATGCAGAGTACATCATGAGAAAGGCTGCACTGGGAAGCAAGCAAGCTAGAAATCAAGATT 13919
QY 1984 GCGGGAGAAAAATAGCAATAACCTCAGATATGAGATGATACCACTCTTATGCGAGAAAGT 2043
Db 13918 GCTGGGAGAAATATCAATAACTCAGTATGAGATGACACCACTCTTATGCGAGAAAGT 13859
QY 2044 GAAGAGAACTAAAAAGCCTCTTGATGAAGGTGAAGAGAGTGAAGAAAGTTGGCTTA 2103

Db 13858 GAAGAGGACTCAAAAGCCTTTGATGAAGTGAAAGAGGAGAGTGGAAAAGTTGGCTTA 13799
 QY 2104 AAGCTCAACATTCAGAAAACGAAGATCATGGCATCTGTCCCATCACTTCATGGAAATA 2163
 Db 13798 AAGCTCAACATTCAGAAAACGAAGATCATGGCATCTGTCCCATCACTTCATGGAAATA 13739
 QY 2164 GATGGGAAAACAGTCGAAACAGTGTACAGACTTTATTTTGGGGGCTCCAAAATCACTG 2223
 Db 13738 GATGGGAAAATGGTGGAAAACAGTGTACAGACTTTATTTTCTGGGCTCCAAAATCACTA 13681
 QY 2224 CAGATGGTGAAGTGCAGCCATGAAATTAAGACACTTACTCTCTGGAGAAAAGTTA--A 2281
 Db 13680 CAGATGGTGAAGTGCAGCCATGAAATTAAGACAGCTTACTCTCTCGAAGGAAAGTTATGA 13621
 QY 2282 CCAACTAGATAGCATATTTGAAGACAGACATTAACCTTCCCAACAAGCCCATCTAG 2341
 Db 13620 CCAACTAGATAGCATATTTGAAGACAAAGACATTAACCTTCCCAACAAGCTCCGCTAG 13561
 QY 2342 TCAAGGCTATGGTTTTCCAGCTGGTGCATGTATGATGTGAGAGTGGAGTCTGGAAGAAAG 2401
 Db 13560 TCAA-----GTTTTCTGTAGTCATGTATGGATGTGAGAGTTGGACTGTGAAGAGG 13508
 QY 2402 CTGAGCATGAAGAAATGATGCTTTTGAACCTGTGTGTGTGGAGAGACTCTTTGAGAGTCC 2461
 Db 13507 CTGAGTCCGGAAGAAATGATGCTTTTGAACCTGTGTGTGTGGAGAGACTCTTTGAGAGTCC 13448
 QY 2462 CTGAGCTGCAAGGAGATCCCAACAGTCCATTCTGAAGGAGATCAGCCCTGGGATTTCTT 2521
 Db 13447 CTGAGCTGCAAGGAGATCCCAACAGTCCATTCTGAAGGAGATCAGCCCTGGGATTTCTT 13388
 QY 2522 TGGAAAGAAATGATGCTAAAGCTGAAATCCCACTTTGGCCACCTCATGTGAAGAGTTG 2581
 Db 13387 TGGAGGAAATGATGCTAAAGCTGAAATCCCACTTTGGCCACCTCATGTGAAGAGTTG 13328
 QY 2582 ACTCACTGGAAGAGACCTGTGCTGGAGGAGTTGGGGCAGGAGGAGAGGAGCGAC 2641
 Db 13327 ACTCACTGGAAGAGACTGTGATGTGGAGGAGTTGGGGCAGGAGGAGAGGAGCGAC 13268
 QY 2642 AGAGATGAGATGGCTGGATGGCATCACTGACTCGATGGAGCTGGGCTGAACTCC 2701
 Db 13267 AGAGATGAGATGGCTGGATGGCATCACTGACTCGATGGAGCTGAGTC--AGTGAATCC 13210
 QY 2702 TGGAGTTGGTATGAGCAGGAGGCGCTGTCTCGGGGATTCATGGGGTCAAAAGATT 2761
 Db 13209 AGGAATTTGGTATGAGCAGGAGGCGCTGGCGTGTGCGATTTCATGGGGTGGCAAAGAGTC 13150
 QY 2762 GGACACGACTGAGCAACTGAACTGAAGTGAAGT 2795
 Db 13149 GGACACGACTGAGTGAAGTGAAGTGAAGT 13116

RESULT 2

ID ABS55700/c
 XX ABS55700 standard; DNA; 38793 BP.

AC ABS55700;

XX 08-JAN-2003 (first entry)

DE Bovine Claudin-16 deficiency associated polynucleotide sequence #3.

XX Bovine; Claudin-16 deficiency; type 2 mutation; ds.

OS Bos sp.

FN JP2002238570-A.

XX 27-AUG-2002.

XX 14-FEB-2001; 2001JLP-00037623.

XX 14-FEB-2001; 2001JLP-00037623.

XX

PA (CHIK-) CHIKUSAN GIJUTSU KYOKAI SH.
 PA (KACH-) KACHIKU KAIRYO JIGYODAN SH.
 XX (DOKU-) DOKURITSU GYOSEI HOJIN KACHIKU KAIRYO CE.
 DR WPI; 2002-744759/81.
 XX
 PT Gene diagnosis of type 2 mutation for bovine claudin-16 deficiency,
 PT comprises checking for mutation in 1-161 position in claudin-16 sequence.
 XX
 PS Example 1; Page 28-41; 43pp; Japanese.
 XX
 CC The invention describes gene diagnosis of type 2 mutation for bovine
 CC claudin-16 deficiency. The method involves (a) getting a bovine nucleic
 CC acid sample; (b) subjecting the sample to a gene amplification reaction
 CC to give a nucleic acid fragment; (c) checking the presence of mutation in
 CC the nucleic acid fragment of step (b). The method is used for gene
 CC diagnosis of type 2 mutation for bovine claudin-16 deficiency. This
 CC sequence represents a polynucleotide from the bovine Claudin-16
 CC deficiency associated gene in which the type 2 mutation occurs
 XX
 SQ Sequence 38793 BP; 11758 A; 7706 C; 7060 G; 12139 T; 0 U; 130 Other;
 Query Match 29.4%; Score 1051.2; DB 1; Length 38793;
 Best Local Similarity 67.4%; Pred. No. 0.03;
 Matches 1496; Conservative 0; Mismatches 123; Indels 602; Gaps 7;
 QY 1187 AAAGATGGAGAGCTCTATACAGTCAGCAAAAACAAGACAGAGCTTACTGTGGCTCAG 1246
 Db 34102 AAAGATGGAGAGCTCTATACAGTCAGCAAAAACAAGACAGAGCTTACTGTGGCTCAG 34043
 QY 1247 ATCATGAACCTCCCTATTGTCMAATTCAGACTTAAATTTGAAGAAAGTAGGCAAAACCACTA 1306
 Db 34042 ATCATGAACCTCCCTATTGTCMAATTCAGACTTAAATTTGAAGAAAGTAGGCAAAACCACTA 33983
 QY 1307 GATCACTCAGGTAGACCTTAATCCAAATCCCTTATGATTATACAGTGGAGTGAAGATA 1366
 Db 33982 GAGCATTCAGGTATGACCTAAATCCAAATCCCTTATGATTATACAGTGGAGTGAAGATA 33923
 QY 1367 GATTTAAGGCGCTTAGATCTGATAGCAGAGTACCTTAATGAACTATGAGCAGAGGTTTCATG 1426
 Db 33922 GATTTAAGGCGCTTAGATCTGATAGCAGAGTACCTTAATGAACTATGAGCAGAGGTTTCATG 33863
 QY 1427 ACATTTGACAGGAGACAGGAGTCCAGACCATCCCATGGAAAAGAAATGCAAAAAGCAA 1486
 Db 33862 ACATTTGACAGGAGACAGGAGTCCAGACCATCCCATGGAAAAGAAATGCAAAAAGCAA 33803
 QY 1487 ATGGCGCTGTGGGGAGCGCTTACAAATAGCTGTG-AAAAGAGAGAACTGAAAAGCAA 1545
 Db 33802 ATGGCGCTGTGGGGAGCGCTTACAAATAGCTGTGAAAAGAGAGAACTGAAAAGCAA 33743
 QY 1546 GGAAAAAGGAAAGATGAAAAGCATCTGAATGCAGAGTTTCCAAA----- 1588
 Db 33742 GGAAAAAGGAAAGATGAAAAGCATCTGAATGCAGAGTTTCAAGAAATAGCAAGAGAGAT 33683
 QY 1589 ----- 1588
 Db 33682 AAGAAAGCCTTCTCAGTGATTCATGCAAGAGATAGAGAAAACAAAGAAATGGAAAG 33623
 QY 1589 ----- 1588
 Db 33622 ACTAGAGATCTCTTCAAGAAAATTAGAGATACCAAGGGGAAAATTTTCATGGAAAAGATGGGA 33563
 QY 1589 ----- 1588
 Db 33562 TCCATAAAGGACAGAGATGGTATGGACTTAACCGGAAGCAGAGATATTAAAGAGAGGTGG 33503
 QY 1589 ----- 1588
 Db 33502 CAGGGATACAGAGAACTGTACAAAARCGAACTTCAAGATCAAGATATACGATGTGGT 33443
 QY 1589 ----- 1588
 Db 33442 TGATCACTCACCTAGAGCAAGATATCCTTGAATGTGAAGTCAATGGGCTTAGAAGCA 33383

QY 1589 ----- 1588
 Db 33392 TCACATGACAAAGCTAGTGGAGGTGATGGAATTCAGTTGAGCTATTTCABAATCTCGA 33323
 QY 1589 ----- 1588
 Db 33322 AAGATGATGCTGTGAAGTGTGCTGCAATTCATATGCCAGCAAAATTTGGAAAATTCAGCAGT 33263
 QY 1589 ----- 1588
 Db 33262 GCCACAGACTGGAAAGGTGAGTTTTCATCCCAATCTCAAGAAAGCAATGCCAAGA 33203
 QY 1589 ----- 1588
 Db 33202 ATGCTCAAACTACCGCACAAATTGMACTCATCTCACACACTAGTAAAGTAATGCTCAAAAT 33143
 QY 1589 -----GAACTTCAGTTGTTCAAGCTGGT 1612
 Db 33142 TCTCCAGCCAGGCTTCAACWAAATATGGGAACGTGTGAACCTGTCAGATGTTTCAAGCTAGT 33083
 QY 1613 TTTAGAAAAGTCAGAGGAACAGAGACCAAAATTCGCAATCTCTCTGTATCATGGA AAAA 1672
 Db 33082 TTTAGAAAAGGCAGAAAACACAGAGACCAANTTCCACATCCGCTGGATCAGCAAAA 33023
 QY 1673 GCAAGAGATTCAGAAAACCAATCTATTCTGCTTTATTGACTATGCAAAAGCCCTTGAC 1732
 Db 33022 GCAAGAGAG-TCCAAAGAAAACATCTATTCTGCTTTATTGACTATGCAAAAGCCCTTGAC 32964
 QY 1733 TGTGGGGTCACAATAAATCTGTGGAATAATCTGAAAGGATGGGAATACACAGACCACTG 1792
 Db 32963 TGTGGGATCAAGAAACTGTGGGAATTTCTGAAGAGATGGGAATACACAGACCACTG 32904
 QY 1793 ACCTGACTCTTGAAAAATTTGTATCAGGTGAGGAAGCAACAGTTAGAACTGGACATGGA 1852
 Db 32903 ACCTGCTCTTGAGAAACCTATCTGAGGTGAGGAAGCAACAGTTAGAACTGGACATGGA 32844
 QY 1853 ACAACAGACTGTTTCCAGTAGGAAGAGATGTCAGAGCTGTATATCTGACCCGGC 1912
 Db 32843 ACAACAGACTGTTTCAACAGGAAGAGATGATCAAGGCTGTATATCTGACCCCTG 32784
 QY 1913 TTGTTTAACTTCTATGACAGACATCATGAGAAACGCTGGCTGGGAAGCAACAGCTG 1972
 Db 32783 TTATTTAACTTATATGACAGATGATCATGAGAAAGCTGGCTGGGAAGCAACAGCTG 32724
 QY 1973 GAATCAAGTTCCCGGAGAAATAGCAATTAACCTCAGATATGCAAGATACCACTTA 2032
 Db 32723 GAATCAAGATTTCTGGGAGAAATATCAATAACCTCAGATATGCAAGATACCACTTTTA 32664
 QY 2033 TGGCAGAAAGTCAAGAGAACTTAAAGGCTCTGATGAAGTGAAGAGAGAGTGA 2092
 Db 32663 TGGCAGAAAGTGAAGGAACTTAAAGGCTCTGATGAAGCAAAAGAGAGAGTGA 32604
 QY 2093 AAGTTGGCTTAAAGCTCAACATTCAGAAACGAGATCATGGCATCTGGTCCCATCACTT 2152
 Db 32603 AATTTGGCTT-AAGCTCAACATTCAGAAACGAGATCATGGCATCTGGTCCCATCACTT 32545
 QY 2153 CATGGGAATAGATGGGGAACAGTGGAAACAGTGTCAAGCTTTATTTTGGGGGGCTC 2212
 Db 32544 CATGGGAATAGATGGGGAACAGTGGAAACAGTGTCAAGCTTTATTTT-CTGGGGCTC 32487
 QY 2213 CAAAATCACTGAGATGCTGACTGAGGCATGAAATTAAGACACTTACTCTCTGGAG 2272
 Db 32486 CAAAATCACTGAGATGCTGACTGAGGCATGAAATTAAGACACTTACTCTCTGGAG 32427
 QY 2273 AAAAGTTA-ACCAACTAGATAGATATGAAAGCAGAGACATTAACCTGCAACAAA 2330
 Db 32426 AAAAGTTATGACCAACCTAGATAGATATTAAGAGCAGAGACATTAACCTGCAACAAA 32367
 QY 2331 GCGCCATCTAGTCAAGGCTATGTTTCTCAGTGGTCTATGATGGATGAGAGTTGGAC 2390
 Db 32366 GGTCCCTCTAGTCAAGGCTATGTTTCTCAGTGGTCTATGATGGATGAGAGTTAGAC 32307

QY 2391 TGTGAAGAAAGCTGAGCACTGAGAAATTTGATGCTTTTGAACCTGCTGCTGGAGAGACT 2450
 Db 32306 TGTGAAGAAAGCTGAGCACTGAGAAATTTGATGCTTTTGAACCTGCTGCTGGAGAGACT 32247
 QY 2451 CTTGAGAGTCCCTTGGACTGCAAGGAGATCCAAACCACTCCATTTCTGAAGAGAGATCAGCCC 2510
 Db 32246 CTTGAGAGTTCCTTGGACTGCAAGGAGATCCAAACCACTCCATTTCTGAAGAGAGATCAGCCC 32187
 QY 2511 TGGGATTTCTTTGGAAGGAATGATCTAAAGCTGAAATCCAGTACTTTTGGCCACTGAT 2570
 Db 32186 TGGGATTTCTTTGGAAGGAATGATCTAAAGCTGAAATCCAGTACTTTTGGCCACTGAT 32127
 QY 2571 CAGAGAGCTGACTCACTGGAAGGACCTGATGCTGGAGGATTTGGGGCAGGAGAG 2630
 Db 32126 GCGAAGAGTTGACCATTTGGAAGAA--CTGATGCTGGAGGATTTGGGGCAGGAGAG 32069
 QY 2631 AAGGGACGACAGAGATGAGATGCTGGATGGCACTGATGCTGATGGAGAGTGTCT 2690
 Db 32068 GAGGGATGACAGAGATGAGATGCTGGATGGCACTGATGCTGATGGAGTGTCT 32009
 QY 2691 GGGTGAACCTCTGGAGTTGGTATGACAGAGGAGGCTGCTGCTGGGATTCATGGGT 2750
 Db 32008 GAATGAACCTCCGGAGTTGATGATGACAGAGGAGGCTGGTGTGCTGATTTCTGGGT 31949
 QY 2751 CACAAAGAGTTGGACACCACTGAGCAACTGAACTGAACTGAACTGAACTGAACTTGT 2810
 Db 31948 TACAAAGAGTCCGACACTACTGAGCGACTGAACTGAACTGAACTGAACTGAACTG 31889
 QY 2811 A 2811
 Db 31888 A 31888
 RESULT 3
 AAD26486/c
 ID AAD26486 standard; DNA; 17141 BP.
 XX
 AC AAD26486;
 DT 26-MAR-2002 (first entry)
 XX Bovine gamma globin gene.
 DE Bovine; xenotransplantation; Gal-alpha (1,3)Gal; Gal determinant;
 KW alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;
 KW surgery; gene therapy; gamma globin gene; ds.
 OS Bos sp.
 XX WO200188096-A2.
 XX 22-NOV-2001.
 XX 14-MAY-2001; 2001WO-US015765.
 XX 15-MAY-2000; 2000US-0204148P.
 XX 13-JUN-2000; 2000US-00593316.
 XX (GERO-) GERON CORP.
 XX Denning C, Clark J;
 XX WPI; 2002-089848/12.
 XX New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal
 PT determinants, for xenotransplantation, and in the treatment of the human
 PT body by surgery or therapy.
 XX Claim 18; Page 70-74; 86pp; English.
 XX The patent discloses immunologically compatible animal tissue, suitable
 CC for xenotransplantation into human patients. The invention also relates
 CC to an ovine tissue devoid of antibody-detectable Gal alpha (1,3)Gal (GAL)

CC determinants which are made by alpha (1,3) galactosyltransferase (GT). The
 CC ovine tissue is useful for treatment of human body by surgery or therapy
 CC and in xenotransplantation, by transplanting the ovine tissue into a
 CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.
 CC Polynucleotide constructs of the invention are useful for inactivating an
 CC alpha(1,3) GT gene in an ovine cell. Alpha (1,3) GT sequences are useful
 CC for drug screening and for the production of GAL containing synthetic
 CC oligosaccharides. Sequences of the invention are also useful in gene
 CC therapy. The present sequence is bovine gamma globin gene

XX SQ Sequence 17141 BP; 5078 A; 3416 C; 3372 G; 5275 T; 0 U; 0 Other;

Query Match 28.8%; Score 1028.7; DB 1; Length 17141;
 Best Local Similarity 91.9%; Pred. No. 0.079;
 Matches 1117; Conservative 0; Mismatches 93; Indels 5; Gaps 3;

QY 1599 GAATTCACGTTGTTCAAGCTGGTTTGTAGAAAAGTCAGAGAACCCAGAGACCAAAATGGCC 1648
 Db 16837 GAACTTCTAGATGTTCAAGCTGGATTAGAAAAGCCAGAGAACCAAAATGGCC 16778

QY 1649 AACATCCTCTGTATCATGGAAGAACAGAGAGTTCCAGAAAACATCTATTTCTGCTTT 1708
 Db 16777 AACATCTGCTGGATCATGGAAGAACAGAGAGTTCCAGAAAAGCATCTATTTCTGCTTT 16718

QY 1709 ATTGACTATGCAAAAGCCTTTGACTGTGGGGGTCACAATAAACTGTGAAAATTTCTGAAA 1768
 Db 16717 ATTGACTATGCCAACCCCTTTGACTGTGGATCACATCACTGTGGAATAATTTCTGAAA 16658

QY 1769 GGGATGGGAATACAGACACCTGACCTGACTCTTTGAAAATTTGTTATGCAGTTCAGGAA 1828
 Db 16657 GAGATGGGAATACCATACCACTGATCTGCTCTTGTAGAAAATTTGTTATGCAGTTCAGGAA 16598

QY 1829 GCAACAGTTAGAACTGGACATGGAACCAACAGACTGTTCCAAAGTAGGAAAAGGATATGT 1888
 Db 16597 GCAACAGTTAGAACTGGACATGGAACCAACAGACTGTTCCAAATAGGAAAAGGATTCGT 16538

QY 1889 CAAGGCTGTATTTGTCACCGGCTGTTTAACTTCTATGCAGAG-ACATCATGAGAAAC 1947
 Db 16537 CAAGGCTGTATTTGTCACCGCTGTTTAACTTCTGTCAGAGTACATCATGAGAAAC 16478

QY 1948 GCTGGGCTGGAAGACACAGCTGGAATCAAGATTCGCCGGGAGAAATAGCAATTAACCTC 2007
 Db 16477 GCTGGACTGGAAGAAACACAGCGCGGAATCAAGATTCGCCGGGAGAAATATCAATAACCTC 16418

QY 2008 AGATATGCAGATGATACCACTTTATGCGAGAAAGTGAAGAGAACTTAAAGGCTCTTG 2067
 Db 16417 AGATATGCAGATGATACCACTTTATGCGAGAAATGAAGGAACTTAAAGGCTCTTG 16358

QY 2068 ATGAAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAACGAAG 2127
 Db 16357 ATGAAGTGAAGTGAAGAGTGAAGAAAGTTGACTTAAAGCTCAACATTCAGAAAATGAAG 16298

QY 2128 ATCATGGCATCTGGTCCCATCCTTCATTCGGAATAGATGGGAAACAGTGGAAACAGTG 2187
 Db 16297 ATCATGGCATCTGGTCCCATCCTTCATTCGGAATAGATGGGAAACAGTGGAAACAGTG 16238

QY 2188 TCAGACTTTATTTTGGGGGCTCCAAAATTCATTCAGATGGTGAATGCAGGCTGAGAA 2247
 Db 16237 GCAGACTTTATTTT--CTGGGCTCCAAAATTCATTCAGATGGTGAATGCAGGCTGAGAA 16180

QY 2248 TTTAAAGACATTTACTCTTGGAGAAAGTTA--ACCACTAGATAGCATATTTGAAA 2305
 Db 16179 TTTAAAGATGTTTACTCTTGGAGAAAGTTATGACCACTAGATAGCATATTTGAAA 16120

QY 2306 GCAGAGACATTACCTTCCCAACAAAGCCCATCTAGTCAAGGCTATGTTTTCAGTGG 2365
 Db 16119 GCAGAGACATTACCTTCCCAACAAAGGTTCTGTAGTCAAGGCTATGTTTTCCTGTGG 16060

QY 2366 TCATGTATGATGTGAGATTGAGCTGTGAGAAAGCTGAGCACTGAAGAAATTTGATGCTT 2425
 Db 16059 TCATGTATGATGTGAGATTGAGCTGTGAGAAAGCTGAGCACTGAGCACTGAGCACTT 16000

QY 2426 TTGAACCTGTGGTGTGGAGAGACTCTTTGAGAGTCCCTTTGGACTGCAAGGAGATCCAAAC 2485

Db 15999 TTGAACCTGTGGTGTGGAGAGACTCTTGAGAGTCCCTTTGGACTGCAAGGAGATCCAAAC 15940
 QY 2486 AGTCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGA 2545
 Db 15939 AGTCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGA 15880
 QY 2546 AACTCCAGTACTTTGGCCACCTGATCAGAAAGCTGACTCCTTGGAAAAAGACCTTGATGC 2605
 Db 15879 AACTCCAGTACTTTGGCCACCTCATGCGAAGAGTTGACTCATTGGAAAAAGACTCTGATGC 15820
 QY 2606 TGGAGGGATTGGGGCAGGAGAGAGGAGGACGACAGAGGATGAGATGGCTGGATGGCA 2665
 Db 15819 TGGAGAGATTGGGGCAGGAGAGAGGAGGACGACAGAGGATGAGATGGCTGGATGGCA 15760
 QY 2666 TCACTGACTCGATGGACGTCGAGTCTGGGTGAACCTCTCGAGTTGGTGTGACACAGGAGG 2725
 Db 15759 TCACTGACTCGATGGACGTCGAGTCTGGGTGAACCTCGGGAGTTGGTGTGACACAGGAGG 15700
 QY 2726 CTTGCTCTCGGGGATTCATGGGGTCAAAAAGTGTGGACAGACTGAGCAACTGAACTG 2785
 Db 15699 CTTGGCGTCTGCAATTCATGGGGTGGCAAGAGTGGGACACGACTGAGCGACTGATCTG 15640
 QY 2786 AACTGAACCTGTACTG 2800
 Db 15639 ATCTGATCTGATCTG 15625

RESULT 4
 ABV99701
 ID ABV99701 standard; cDNA; 78056 BP.

XX AC ABV99701;
 XX XX
 DT 12-FEB-2003 (first entry)
 XX DE Bovine BSE-resistant prion protein cDNA.
 DE DE
 KW Transmissible spongiform encephalopathy; neuroprotective; prion protein;
 KW bovine spongiform encephalopathy; transgenic; BSE; bovine; cervid; PrP;
 KW TSE; gene; ss.
 XX OS Bos taurus.
 OS OS
 PN WO200279416-A2.
 PN XX
 PD 10-OCT-2002.
 PD XX
 PF 28-MAR-2002; 2002WO-US009652.
 PF XX
 PR 30-MAR-2001; 2001US-0280549P.
 PR XX

XX (TEXA) UNIV TEXAS A & M SYSTEM.

XX Dunne PW, Piedrahita J;

XX WPI; 2003-092895/08.

DR P-PSDB; ABP57900.

XX New transgenic bovine and cervid useful for producing animals which are
 PT resistant to bovine spongiform encephalopathy and transmissible
 PT spongiform encephalopathy disease, comprise a transgene encoding a mutant
 PT PrP polypeptide.

XX Disclosure; Page 60-82; 98pp; English.

XX The invention relates to a novel transgenic bovine/cervid comprising a
 CC transgene encoding a mutant prion protein (PrP) polypeptide, in which a
 CC substitution has been made at position 171 of the sequence, which renders
 CC the bovine/cervid resistant to bovine spongiform encephalopathy (BSE) and
 CC transmissible spongiform encephalopathy (TSE) disease, respectively. The
 CC transgene of the invention has neuroprotective activity. The method is
 CC useful for producing a transgenic bovine or cervid resistant to BSE and

CC TSE diseases. The bovine prion gene is useful for producing transgenic
 CC cattle exhibiting resistance to bovine spongiform encephalopathy. The
 CC sequence represents the transgene encoding the mutant bovine PrP
 CC polypeptide
 XX
 SQ Sequence 78056 BP; 22052 A; 17462 C; 16284 G; 22258 T; 0 U; 0 Other;
 Query Match 28.3%; Score 1012.1; DB 1; Length 78056;
 Best Local Similarity 92.8%; Pred. No. 0.02;
 Matches 1113; Conservative 0; Mismatches 79; Indels 7; Gaps 5;
 1589 GAACCTTCAGTTGTTCAAGCTGTTTATAGAAAGTCAGAGAACCCAGAGACCAAAATGCCC 1648
 33273 GAACCTTCAGTTGTTCAAGCTGTTTATAGAAAGTCAGAGAACCCAGAGATCAAAATGCC 33332
 1649 AACATCCTCTGATCATGGAAGAAAGCAAGAGAGTTCCAGAAACATCTATTTCTGCTTT 1708
 33333 AACATCCTCTGATCATGGAAGAAAGCAAGAGTTCCAGAAACATCTATTTCTACTTT 33392
 1709 ATTCACTATGCAAGAGCTTTGACTGTGGGGTCACATATACTGTGGAAATTCGAAA 1768
 33393 ATTCACTATGCAAGAGCTTTGACTGTGGGGTCACATATACTGTGGAAATTCGAAA 33452
 1769 GGGATGGGAATPACAGAGCCCTGACCTGATCTTGTGAAAAATTTGTATGAGGATGATGT 1828
 33453 GAGATGGGAATPACAGAGCCCTGATCTGCTCTTGTGAAAAATTTATATGAGGATGATGT 33512
 1829 GCAACAGTTAGAACTGGACATGGAACACAGAGCTGTTCCAGTAGGAAAGAGATGT 1888
 33513 GCAACAGTTAGAACTGGACATGGAACACAGAGCTGTTCCAAATAGGAAAGAGATGT 33572
 1889 CAAGCTCTATATTGTACACCGGCTTGTAACTTCTATGAGAG-ACATCATGAGAAAC 1947
 33573 CAAGCTCTATATTGTACACCGGCTTGTAACTTCTATGAGAGTACATCATGAGAAAC 33632
 1948 CTGGCTGGGAAGAACCAAGCTGGAATCAAGATTGCGGGAGAAATAGCATTAACCTC 2007
 33633 CTGGCTGGGAAGAACCAAGCTGGAATCAAGATTGCGGGAGAAATATCAATAACCTC 33692
 2008 AGATATGAGATGATACACACCTTATGCGAGAAAGTGAAGAGAACTTAAAGAGCTCTTG 2067
 33693 AGATATGAGATGATACACACCTTATGCGAGAAAGTGAAGAGAACTTAAAGAGCTCTTG 33752
 2068 ATGAAGGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAG 2127
 33753 AGGAAGGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAG 33812
 2128 ATCATGGCATCTGCTCCCATCACTTCATGGAAATAGATGGGAAACAGTGGAAACAGTG 2187
 33813 ATCATGGCATCTGCTCCCATCACTTCATGGAAATAGATGGGAAACAGTGGAAACAGTG 33872
 2188 TCAGACTTTATTTTGTGGGGGCTCCAAATCACTGAGATGTTGACTGAGCCATGAAA 2247
 33873 TCAGACTTTATTTT-CTGGGCTCCAAATCACTGAGATGTTGACTGAGCCATGAAA 33930
 2248 TTAAGAAGCACTTACTCTCTGGAGAGAAAGTTA--ACCACTTAGATAGCATTTGAAA 2305
 33931 TTAAGAAGCACTTCTCTCTGGAGAGAAAGTTATGACCACTTAGATAGCATTTGAAA 33990
 2306 GCAGAGCATTAACCTTGGCAACAAAGCCCTTGTGAGTCAAGCTTATGTTTTCAGTGG 2365
 33991 -CAGAGAGCTTACTTTGGCAACAAAGTTCTGCTTAGTCAAGCTATGG-ITTTCTGTGG 34048
 2366 TCATGATGAGTGTGAGAGTTGGAATGTGAAGAAAGCTGAGCACTGAAGAAATGATGCTT 2425
 34049 TCATGATGAGTGTGAGAGTTGGAATGTGAAGAAAGCTGAGCACTGAAGAAATGATGCTT 34108
 2426 TTGAAGCTGCTGTTGGAGAGAGCTTCTGAGAGTCCCTTGGAGTCAAGGAGATCCAAAC 2485
 34109 TTGAAGCTGCTGTTGGAGAGAGCTTCTGAGAGTCCCTTGGAGTCAAGGAGATCCAAAC 34168
 2486 AGTCCATTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTGA 2545

Db 34169 AGTCCATCTGAAGGACATCAGCCCTGGGATTTCTTTTGAAGGAATGATGCTAAAGCTGA 34228
 QY 2546 AACTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCACTGAAAAGACCTGATGC 2605
 Db 34229 AACTCCAGTACTTTGGCCACCTGATGAGAGTTGACTTCTGAAAAGACTCTGATGC 34288
 QY 2606 TGGGAGGATTTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2665
 Db 34289 TGGGAGGATTTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 34348
 QY 2666 TCATGACTTCGATGAGACGCTGAGTGTGGTGAATCTCTGAGAGTTGATGAGACAGGAGG 2725
 Db 34349 TCATGACTTCGATGAGTGTGAGTGAATCTCTGAGTGAATCTCTGAGTGAATCTCTGAGGAGG 34408
 QY 2726 CCTGTCTTGGCGGATTCATGGGGTCAAAAAGAGTTGGACACGACTGAGCAACTGAACT 2784
 Db 34409 CCTGGCTGTGTGATTCATGGGGTCCAAAAGTCCGAGCTGAGAGACTGATCT 34467
 RESULT 5
 AAZ58843/C
 ID AAZ58843 standard; DNA; 10492 BP.
 XX
 AC AAZ58843;
 XX
 DT 25-APR-2000 (first entry)
 XX
 DE Nucleotide sequence of bovine myostatin promoter-enhancer region.
 XX
 KW Myostatin; promoter; muscle cell disorder; genetic marker; bovine;
 KW myogenic regulatory factor; oncogene; muscle growth;
 KW muscular dystrophy gene; ss.
 XX
 OS Bos sp.
 XX
 PH Key Location/Qualifiers
 enhancer 5094..9099
 FT /tag= a
 FT /note= "E-box 4"
 FT 9236..9241
 FT /tag= b
 FT /note= "E-box 3"
 FT 9860..9865
 FT /tag= c
 FT /note= "E-box 2"
 FT 9879..9885
 FT /tag= d
 FT /note= "TATA-box 3"
 FT 10095..10100
 FT /tag= e
 FT /note= "E-box 1"
 FT 10197..10201
 FT /tag= f
 FT /note= "CAAT-box"
 FT 10240..10246
 FT /tag= g
 FT /note= "TATA-box 2"
 FT 10264..10270
 FT /tag= h
 FT /note= "TATA-box 1"
 FT 10403..10492
 FT /tag= i
 FT /note= "partial coding fragment"
 XX
 PN WO200001810-A1.
 XX
 PD 13-JAN-2000.
 XX
 PF 07-JUL-1999; 99WO-NZ000107.
 XX
 PR 07-JUL-1998; 98NZ-00330902.
 XX
 PA (NZPA-) NEW ZEALAND PASTORAL AGRIC RES INST LTD.

XX Bass JJ, Jeanplong F, Kambadur R, Sharma M;
 XX WP1; 2000-182112/16.
 XX Novel promoter sequences of the myostatin gene used to diagnose muscle
 XX cell disorders.
 XX Claim 1; Fig 1; 32pp; English.
 XX The invention relates to a bovine myostatin promoter sequence. The
 CC myostatin promoter polynucleotide sequences can be used in diagnostic
 CC tests for muscle cell disorders, and as a genetic marker for the
 CC selection, in particular, of cattle and sheep breeds that express low
 CC levels of myostatin, and so have increased muscle mass. The promoter can
 CC also be used for the transcriptional control of a heterologous gene of
 CC interest, e.g. myogenic regulatory factors, myostatin and myostatin
 CC receptor, oncogenes, genes that regulate muscle differentiation and
 CC growth, the muscular dystrophy gene, and any other genes expressed in
 CC muscle. The present sequence represents the nucleotide sequence of bovine
 CC myostatin promoter-enhancer region
 XX
 SQ Sequence 10492 BP; 3262 A; 2160 C; 2002 G; 3068 T; 0 U; 0 Other;
 Query Match 28.2%; Score 1009; DB 1; Length 10492;
 Best Local Similarity 91.9%; Pred No. 0.15; Mismatches 85; Indels 14; Gaps 6;
 Matches 1127; Conservative 0;
 QY 1585 CAAGAATCTTCAGTGTTCAGTGTGTTTGTAGAAAGTCAGAGAACACAGACCAAT 1644
 DB 5545 CCATGAATCTTCAGTGTTCAGTGTGTTTGTAGAAAGTCAGAGAACACAGATCAAT 5486
 QY 1645 TGCCAAATCTCTGTATCATGGAAGAGAGAGAGTCCAGAAACATCTATTTCGT 1704
 DB 5485 TGCCAAATCTCTGTATCATGGAAGAGAGAGTCCAGAAACATCTATTTCGT 5426
 QY 1705 CTTTATTGACATGCAAGGCTTTGACTGTGGGGTCACATATAACTGTGGAAATCT 1764
 DB 5425 CTTTATTGACATGCAAGGCTTTGACTGTGGATCACAATAAATGTCTCAAAATCT 5366
 QY 1765 GAAAGGATG-----GGATACACAGACACCTGACCTGTTGAAAATTTGTAT 1816
 DB 5365 GAAAGAGATGTTCTGACCACTGACCTGCCACCTGACCTGCTCTTGAGAAACCTGTAT 5306
 QY 1817 GCAGTCAGGAAGCAACAGTTAGAACTGCATGCAACAGAGCTGTTCCAGTAGGA 1876
 DB 5305 GCAGTCAGGAAGCAACAGTTAGAACTGCATGCAACAGAGCTGTTCCAAATAGGA 5246
 QY 1877 AAGAGATGATCAAGGCTGTATATTGTCACCCGGCTTGTAACTTCTATGCAGAG-AC 1935
 DB 5245 AAGAGAGATGTCAGAGCTGTATATTGTCACCCCTGCTTATTTAACTTCTATGCAGAGTAC 5186
 QY 1936 ATCATGAAACGCTGGCTGGAGAGACAGCAAGCTGGAATCAGATGTCGGGAGAAAT 1995
 DB 5185 ATCATGAAACACTGGCTGGAGAGACAGCAAGCTGGAATCAGATGTCGGGAGAAAT 5126
 QY 1996 AGCAATACCTCAGATATGAGATGATACCACTTATGGCAGAAAGTGAAGAGAACTA 2055
 DB 5125 ATTAATACCTCAGATATGAGATGATACCACTTATGGCAGAAAGTGAAGAGAACTC 5066
 QY 2056 AAAAGCCTCTTGATGAAGTGAAGAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACT 2115
 DB 5065 AAAAGCCTCTTGATGAAGTGAAGAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACT 5006
 QY 2116 CAGAAACGGAAGATCATGGCTCTGTCCTCATCTTCTATGGGAAAT-AGATGGGAAAC 2174
 DB 5005 CAGAAACGGAAGATCATGGCTCTGTCCTCATCTTCTATGGGAAAT-AGATGGGAAAC 4946
 QY 2175 AGTGAAACAGTGTGACATTTATTTTGGGGGGCTCCAAATCACTGCAGATGTTGAC 2234
 DB 4945 AGTGAAACAGTGTGAC-AGACTTTATTTTCTGGGCTCCAAATCACTGCAATTTGTTGAC 4887
 QY 2235 TGCAGCCATGAATTAAGAGACACTTACTCTCTTGGAGAAAGTTA--ACCAACCTAGAT 2292

DB 4886 TGCAGCCATGAATTAAGAGATGCTTACTCTCTTGGAGAAAGTTATGACCAACCTAGAT 4827
 QY 2293 AGCATATTGAAAGACAGACATTAACCTTCCCAACAAAGCCCATCTAGTCAAGGCTATG 2352
 DB 4826 AGCATATTGAAAGACAGACATTAACCTTCCCAACAAAGCTCCATCTAGTCAAGGCTATG 4767
 QY 2353 GTTTTCCAGTGTTCATGTATGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCACTGA 2412
 DB 4766 GTTTTCCAGTGTTCATGTATGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCCGA 4707
 QY 2413 AGAATTGATCTTTTGAATCTGTGGTGTGAGAGAACTCTTGAAGTCCCTTGGACTGCA 2472
 DB 4706 AGAATTGATCTTTTGAATCTGTGGTGTGAGAGAACTCTTGAAGTCCCTTGGACTGCA 4647
 QY 2473 AGAGATCCAAACAGTCCATCTTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATG 2532
 DB 4646 AGAGATCCAAACAGTCCATCTTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATG 4587
 QY 2533 ATGCTAAAGCTGAATCCAGTACTTTGGCCACTGATCAGAGAGCTGACTCAGCTGAA 2592
 DB 4586 ATGCTAAAGCTGAATCCAGTACTTTGGCCACTGATCAGAGAAAGTTGACTCATTTGAA 4527
 QY 2593 AAGACCTGATCTGGAGGGATTGGGGCAGAGAGAGAGAGGAGGAGGAGGAGATGAGA 2652
 DB 4526 AAGACTCTGATCTGGAGGGATTGGGGCAGAGAGAGAGAGGAGGAGGAGATGAGA 4467
 QY 2653 TGCTGTGATGGCATCACTGACTCGATGGAGCTGAGTCTGGGTGAACTCTCTGAGTGGTG 2712
 DB 4466 TGCTGTGATGGCATCACTGACTCGATGGAGCTGAGTCTGGGTGAACTCTCTGAGTGGTG 4407
 QY 2713 ATGGA-CAGGAGGCTGTCTCGGGGATTCATGGGGTCAAAAGATTGGAACAGT 2771
 DB 4406 ATGGA-CAGGAGGCTGTCTCGGGGATTCATGGGGTCAAAAGATTGGAACAGT 4347
 QY 2772 GAGCAACTGAACCTGAACCTGTA 2797
 DB 4346 GAAGGACTGAACCTGAACCTGTA 4321
 RESULT 6
 ABK14793/c
 ID ABK14793 standard; DNA; 23653 Bp.
 XX AC ABK14793;
 XX DT 08-MAY-2002 (first entry)
 XX DE Platelet-derived growth factor expression cassette insert used in pBC734.
 XX KW Platelet-derived growth factor; PDGF; transgenic; milk; gene;
 XX KW wound healing; diabetic foot ulcer; decubitus ulcer; ophthalmic disease;
 XX KW venous stasis ulcer; periodontal regeneration; bone formation;
 XX KW prosthetic vascular graft; pBC734; ds.
 XX OS Homo sapiens.
 XX PN WO200198520-A1.
 XX PD 27-DEC-2001.
 XX PF 19-JUN-2001; 2001WO-US041044.
 XX PR 19-JUN-2000; 2000US-0212406P.
 XX DA (GENZ) GENZYME TRANSGENICS CORP.
 XX PI Echelard Y, Meade H, Eichner W, Sommermeyer K;
 XX DR WPI; 2002-083329/11.
 XX PT Production of platelet derived growth factor (PDGF) comprises expression
 in the milk of a non-human transgenic animal.

PN WO200279416-A2.
XX 10-OCT-2002.
XX 28-MAR-2002; 2002WO-US009652.
XX 30-MAR-2001; 2001US-0280549P.
XX (TEXA) UNIV TEXAS A & M SYSTEM.
XX Dunne PW, Piedrahita J;
XX WPI, 2003-092895/08.
XX P-PsDB; ABP57900.
XX New transgenic bovine and cervid useful for producing animals which are
PT resistant to bovine spongiform encephalopathy and transmissible
PT prioniform encephalopathy disease, comprise a transgene encoding a mutant
PT PrP polypeptide.
XX Disclosure; Page 60-82; 98pp; English.
XX The invention relates to a novel transgenic bovine/cervid comprising a
CC transgene encoding a mutant prion protein (PrP) polypeptide, in which a
CC substitution has been made at position 171 of the sequence, which renders
CC the bovine/cervid resistant to bovine spongiform encephalopathy (BSE) and
CC transmissible spongiform encephalopathy (TSE) disease, respectively. The
CC transgene of the invention has neuroprotective activity. The method is
CC useful for producing a transgenic bovine or cervid resistant to BSE and
CC TSE diseases. The bovine prion gene is useful for producing transgenic
CC cattle exhibiting resistance to bovine spongiform encephalopathy. The
CC sequence represents the transgene encoding the mutant bovine PrP
CC polypeptide
XX
SQ Sequence 78056 BP; 22052 A; 17462 C; 16284 G; 22258 T; 0 U; 0 Other;
Query Match 24.8%; Score 885.401; DB 1; Length 78056;
Best Local Similarity 86.6%; Pred. No. 0.053;
Matches 1063; Conservative 0; Mismatches 151; Indels 14; Gaps 8;
QY 1573 AATGACAGTCCCAAGAACCTCCAGTGTTCAGCTGGTTTATAGAAAGTCAGAGAAC 1632
DB 1660 AAGTACATGACCAATGACCTCCAGATGTTTCAAGCTGGATTTAGAAAGGACAGGAAT 1601
QY 1633 CAGACACAAATTTGCAACATCCTCTGTATCATGGAAGAAAGCAAGAGAGTTCCAGAAAA 1692
DB 1600 CAGAGATCAATTTGCCAATCCTATGATCATGGAAGAAAGTAAGAGCGTTCCAGAAAA 1541
QY 1693 CATCTATTCTGCTTTATGACTATGCAAAAGCCTTTGACTGTGGGGTCAATTAACCT 1752
DB 1540 CATCCATTTTGTCTTTATGACTGTGCCAAAGCCTTTGACTGTGTGGATCAACAAACT 1481
QY 1753 GTGGAAATTTCTGAAGGATGGGAATACCAGACCACTGACCTGACTCTTTGAAAAATTT 1812
DB 1480 GTGGGAATTTCTTAAGAGATGGGAATACCAGAGCGCTGACCTGCCCTCTGAGAACTCT 1421
QY 1813 GTATCAGCTCAGGAAGCAACAGTTAGAACTGGAATGGAACACAGACTGGTTCCAAAGT 1872
DB 1420 GTATGCAAGTTAGGAAGCAACAGTTAGAACTAGACATGGAACACATACTGGTTCCAAAT 1361
QY 1873 AGGAAGAGGATATCTCAAGGCTGTATATGTCACCCGCTGTTTAACTTCTATGCGA 1932
DB 1360 CAGGAAGAGGATACATGAGGCTGTATATGTCACCCCTGTTTAACTTATATGCGA 1301
QY 1933 G-ACATCATGGAACCGTGGCTGGAGAGAGCAACAGCTGGAATCAAGATTGCGGGAG 1991
DB 1300 GTACATCAGGAAATGCTGGACTGGATTAAGCAACAGCTGGAATTAAGTTTGGCAGGAG 1241
QY 1992 AATATGCAATACCTCAGATATGCAATGATACACCCCTTATGCGAGAAAGTGAAGAGA 2051
DB 1240 AATATATCAATACCTCAGATATGCAATGATACACCCCTTATGCG--TTAAGAGAGAGA 1183
QY 2052 ACTAAAAAGCCTCTTGTATGAAGCTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAA 2111

DB 1182 CCTAAAGAGCCTTTGATGAAAGTGAAGAGAGAGTGAAGAAAGTTGACTTAAACTCAA 1123
QY 2112 CATTCAGAAAACGAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAATAGATGGGA 2171
DB 1122 CATTCAGAAAACGAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAATAGATGGGA 1063
QY 2172 AACAGTGGAAACAGTGTCAAGCTTTATTTTGGGGGGCTCCAAAATCACTGCAGATGT 2231
DB 1062 AACAGTGGAAACAGTGTCAAGCTTTATTTTGGGGGGCTCCAAAATCACTGCATATGT 1006
QY 2232 GACTGCAGGCATGAATTAATAAGACACTTACTCTTTGGAAGAAAAGTTA--ACCAACCTA 2289
DB 1005 GACTGCAGGCATGAATTAATAAGTCTTCTTGGGAAAGAAAGTTATGACCAACCTA 947
QY 2290 GATAGCATATTAAGACGAGACATTACTTGCACCAACAAAGCCCATCTAGCTCAAGCT 2349
DB 946 GACAGCATATTAATAAGACGAGACATTACTTGCACCAACAAAGCTCCATCTAGTCAAAGCT 887
QY 2350 ATGTTTTTCCAGTGGTGTATGATGTGAGAGTTGAGACTGTGAAGAAAGCTGAGCAC 2409
DB 886 ATGTTTTTCCAGTGTATGATGTGAGAGTTGAGACTGTGAAGAAAGCTGAGTGC 827
QY 2410 TGAAGATGATGCTTTTGAACCTGTGGTGTGGAGAAAGCTCTTGAGAGTCCCTTGGACT 2469
DB 826 TGAAGATGATGCTTTTGAACCTGTGGTGTGGAGAAAGCTCTTGAGAGTCTATTTGGACT 767
QY 2470 GCAAGAGATCCCAACAGTCCATCTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGA 2529
DB 766 GCA---AGATCAACAGCTCCATCTTAAAGGAAATCAGTCCCTGAATATTCATCGGAAGA 710
QY 2530 ATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACTTCAATCAAGAGAGCTGACTCACTG 2589
DB 709 CTAAATGCTGAAGCTGAAGTCCAAATCTTTGGCCACTGATGCAAGAAAGCTGCTCATTTG 650
QY 2590 GAAAAGCCTGTATGCTGGAGGAGTGGGGCAGGAGAGAGGAGGAGGAGGAGGAGGAGATG 2649
DB 649 GAAAAGCCTGTATGCTGGAGATGATGAAGTGGGAGGAGGAGGAGGAGGAGGAGATG 590
QY 2650 AGATGGTGGATGGCATCACTGATGATGGAGCTGAGTCTGGGTGAACCTCTTGGAGTTG 2709
DB 589 AGATGGTGGATGGCATCACTGATGATGGAGCTCAATGAGATGAGTTGAGTAAAGCTCTAGGAGTTG 530
QY 2710 GTATGGAGC-AGGAGGCTGTCTCTGGGGGATTCATGGGGTCAACAAAGAGTTGGACAG 2768
DB 529 GTATGGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 471
QY 2769 ACTGAGCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTG 2796
DB 470 ACTGAGCAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTG 443
RESULT 10
AAQ79534/C
ID AAQ79534 standard; DNA; 3370 BP.
XX AAQ79534;
AC AAQ79534;
XX 25-MAR-2003 (revised)
DT 04-AUG-1995 (first entry)
XX Bovine tracheal antimicrobial peptide gene.
DE Antimicrobial peptide; TAP; ss.
XX Bos taurus.
XX Key Location/Qualifiers
FT intron 1..1472
FT CAAT_signal /*tag= e
FT TATA_signal /*tag= a
FT 1444..1449

FT exon /tag= b
 FT 1503..1560
 FT /tag= c
 FT intron 1561..3073
 FT /tag= f
 FT exon 3074..3211
 FT /tag= d
 FT intron 3212..3370
 FT /tag= g
 XX
 XX WO9426106-Al.
 XX
 XX 24-NOV-1994.
 XX
 XX 11-MAY-1994; 94WO-US005257.
 XX
 XX 11-MAY-1993; 93US-00060822.
 XX
 XX (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
 XX
 XX Zaslloff MA, Bevins CL, Diamond G;
 PI
 XX WPI; 1995-006233/01.
 DR P-PSDB; AAR66205.
 XX
 XX New antimicrobial peptide and precursor from mammalian trachea - and
 PT related DNA, vector and transformed cells, useful as pharmaceuticals and
 PT disinfectants.
 XX
 XX Claim 12; Page 40-42; 69pp; English.
 XX
 XX The tracheal antimicrobial peptide (TAP) gene is expressed primarily in
 CC the columnar epithelial cells of the mucosa lining the bovine airway. RNA
 CC analysis indicates that TAP is expressed along the entire length of the
 CC conducting airway. It includes a consensus sequence for an NFkB
 CC recognition site 181 bases upstream from the transcriptional start site.
 CC When assayed in vitro several different strains of microbes, including
 CC some which are respiratory pathogens, a mammalian TAP (see AAR66204 FT)
 CC of bovine origin showed similar inhibitory activity to that of synthetic
 CC marainin 2-NH₂. While TAP (see AAR66204 FT) was most active against E.
 CC coli and K. pneumoniae, significant antimicrobial activity was also seen
 CC when applied to C. albicans. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX
 XX Sequence 3370 BP; 826 A; 872 C; 811 G; 861 T; 0 U; 0 Other;
 SQ

Query Match 18.8%; Score 672; DB 1; Length 3370;
 Best Local Similarity 89.1%; Pred. No. 6.1;
 Matches 778; Conservative 0; Mismatches 85; Indels 10; Gaps 5;

QY 1950 TGGGCTGGAAGACACCAAGCTGGAAATCAAGATTCCCGGAGAAATAGCAATACCTCAG 2009
 DB 939 TTGGCTGGAAGACACCAAGCTGGAAATCAAGATTCCCGGAGAAATAGCAATACCTCAG 880
 QY 2010 ATATCGCATGATACCACTTTATGGCAGAAAGTGAAGAGAACTAAAAGCCCTTCATG 2069
 DB 879 ATATCGCATGATACCACTTTATGGCAGAAAGTGAAGAGAACTAAAAGCCCTTCATG 820
 QY 2070 GAAAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAACGAAGAT 2129
 DB 819 GAAAGTGAAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAACGAAGAT 760
 QY 2130 CATGCGATCTGGTCCCATCATTCATGCGAAATAGATGGGAAACAGTGGAAACAGTGTC 2189
 DB 759 AATGGCGTATGATCCCATCATTCATGCGAAATAGATGGGAAACAGTGGAAACAGTGTC 700
 QY 2190 AGACTTTATTTTGGGGGGCTCAAAATCACTCGCATGGTGAAGTGCAGCCATGAAT 2249
 DB 699 AGACTTTATTTTGGGGGGCTCAAAATCACTCGCATGGTGAAGTGCAGCCATGAAT 643
 QY 2250 AAAAGACACTTACTCTCTTGGAGAAAGTTA--ACCAACCTAGATAGCATATTGAAAAGC 2307
 DB 642 AAAACACGATTACTCTCTTGGAGAAAGTTAATGACCAACCTAGATAGCAT--TTGAAAAGC 584

QY 2308 AGAGACATTACCTTGGCAACAAGCCCATCTAGTCAAGGC---TATGGTTTTTCCAGTG 2364
 DB 583 AGAGATATTACTTTACCACAAAGGTCGCGCTAGTCAAGGCCTATGGTTTTTCCAGTG 524
 QY 2365 GTCATGATGGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCACTGAAGAAATTGATGCT 2424
 DB 523 GTCATGATGGATGTGAGAGTTGGACTGTGAAG--AAGCTGAGCGCCAAAGAAATTGATGCT 465
 QY 2425 TTTCAACTGTGGTGTGGAGAAAGACTCTTGAAGAGTCCCTTGGACTCAAGGAGATCCAAAC 2484
 DB 464 TTTGAACGTGGTGTGGAGAAAGACTCTTGAAGAGTCTTTTGAAGTCTTGGACTCAAGGAGATCCAAAC 405
 QY 2485 CAGTCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAAGCTG 2544
 DB 404 CAGTCCATCCTAAAGGAGATCAGTCTGGGTGTTCTTTGGAAGGACTCATGCTGAAGCTG 345
 QY 2545 AACTCCAGTACTTTGGCCACCTGATCAGAAGAGCTGACTCACTGGAAGAGACCTGATG 2604
 DB 344 AAGCTCAATACCTTGGCCACCTCATGCGAAGAGTTGACTCATTTGGAAGAGACCTTGGT 285
 QY 2605 CTGGAGGGATTGGGGCAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2664
 DB 284 CTGGAGGGATTGGGGCAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 225
 QY 2665 ATCAGTACTGATGAGAGCTGAGTCTGGGTGAACTCTCTGGAGTTGGTGTGACAGGGAG 2724
 DB 224 ATCAGGACTCGATGAGAGCTGAGTTGAGTGAAGTCTGGGAGTCTGGTGTGAGAGGGAG 165
 QY 2725 GCCTGTCTGGCGGATTTCATGGGGTCAAAAAGATTGGACACGACTGAGCAACTGAACT 2784
 DB 164 TCCTGGCATGCTGCAATTCATGGGTGCAAGAGTTGGTACCCTGAGCGACTGAACT 105
 QY 2785 GAACTGAAGTACTGAAACCTTAGTGGTTAT 2817
 DB 104 GAACTGAAGTACTGAACTGATGGTTTCTTTT 72

RESULT 11
 AAT96051
 ID AAT96051 standard; DNA; 2830 BP.
 AC
 AC AAT96051;
 XX
 XX 01-MAY-1998 (first entry)
 XX
 XX DNA for bovine brain protein p97.
 XX
 XX Bovine; brain protein; p97; genetic engineering; analysis; ds.
 XX
 XX Bos taurus.
 PH Key Location/Qualifiers
 FT CDS 123..1901
 FT /tag= a
 FT /product= "p97"
 XX
 XX JP10001497-A.
 PN
 XX
 XX 06-JAN-1998.
 XX
 XX 13-JUN-1996; 96JP-00152684.
 XX
 XX 13-JUN-1996; 96JP-00152684.
 XX
 XX (MITU) MITSUBISHI CHEM CORP.
 XX
 XX WPI; 1998-114806/11.
 DR P-PSDB; AAW38337.
 XX
 XX New protein and a DNA encoding it - useful in genetic engineering for
 PT various analyses.
 XX

PS Claim 3; Page 6-8; 12pp; Japanese.

XX The present sequence encodes the bovine brain protein p37, which is
 CC useful in genetic engineering techniques for carrying out various
 CC analyses

XX Sequence 2830 BP; 938 A; 522 C; 686 G; 684 T; 0 U; 0 Other;

Query Match 18.7%; Score 567.9; DB 1; Length 2830;
 Best Local Similarity 81.7%; Pred. No. 7.4;
 Matches 782; Conservative 0; Mismatches 166; Indels 9; Gaps 1;

Qy	369	AAGC	AAAAATGATAGGATCTAGAAAGAGGAACTCCCGCAGGTCACTAGTGCCTCCCATATGC	428
Db	659	AAGG	CTCTGGCAGGACATCGAAAGATAGCCCGCAGGTCCGAAAGGGTTTCAACATGC	718
Qy	429	TACTGGAGATCAGTGGAGA-----AA	TAACTCCAGTAAGAATGAAGCGGTGGCA	479
Db	719	TACTGGAGAGAGCGGAGCGCGGATACTTAA	TACCTCCAGTAAGAATGAAGCGGTGGCA	778
Qy	480	AAAGCAAAAGAAATACCCAGCTGTGGATGTGAC	TGTGATATAGCAAGTCCGATCGT	539
Db	779	AAAGTGAAGAGGACAGTCAGCTGCGATGTGCT	TGGTACGAAAGTAACTCCGATCGT	838
Qy	540	TAAAGACGAATATTTCATAGGAACCTGGAATGT	CAGGTCCATGAATCAAGGCAAAATTGGA	599
Db	839	TAAAGAGAAATACTGCATAGGAACCTGGAA	TGTTAGATCTATGAATCCTGGTAAATTGGA	898
Qy	600	AGTGGTCAACAAGAGATGGCAAGAGTGAATGT	CAACATTTAGGAATCAGCGAACTAAA	659
Db	899	TGTGGTGAAGCAGGAGATGGAAGAAATAAC	CATCGATCTTAGGAATCAGTGAACATAA	958
Qy	660	ATGACATGGAATGGTGAATTTAACTCAGATGAC	ATTATATCTACTACTCGGGGACGGA	719
Db	959	ATGACAGGAATGGCGAATTTGAATTCAGATGAC	ATTAATCTATTACTGTGGGCAACA	1018
Qy	720	ATCCCTCAGAGAAATAGGAGTAGCCATCATGGT	CAACAAAAGATCCGAAATGCAGTACT	779
Db	1019	ATCCCTTAGAGAAATAGGAGTCGCTCATAGT	TTAAGAAAAGATCCGAAATGCAATAAT	1078
Qy	780	TGGATCGAGTCTCAAAACGACGAATGATCTCT	GTGTGTTTCCAAAGCAAAACCATTCAA	839
Db	1079	TGGGTGCATCTGAAAACGACGAGATGATTTCA	GTGTTTCCAAAGCAAAACCATTCAA	1138
Qy	840	TATCACAGTAAATCCAAGTCTATGCCCAACCA	AGTAATGCTGAAGAAGCTGAAGTTGAACG	899
Db	1139	CCTCACAGTAAATCCAAGTCTATGCCCAACTC	CTTATGCTGAAGAAGCTGAAGTTTACCG	1198
Qy	900	GTCCTATGAAGACCTACAGACCTTTTAGACTA	ACACCCCAAAAAGATGTCTCTTCAT	959
Db	1199	GTTCTATGAAGACTACACACCTTCTGGAATAT	AACCCGAAAATAGATGTCTTTTCAT	1258
Qy	960	TATAGGGACCTGGAATGCAAAGTAGGAAGCAA	AGAAACACCTCGAGTAACGAGCAAAAT	1019
Db	1259	CATAGGGGATTTGGAAATGCAAAGATGGGAAG	TCAAGAGATACCTGGAAATACAGGACGTT	1318
Qy	1020	TGGCCTTGGAAATACGGAATGAAGCAGGGCA	AAAGACTAATAGATTTTGGCCAAAGAAATGC	1079
Db	1319	TGGCCTTGGAAATGCAAAATGAAGCAGGGCGA	AGGCTAATCAGATTTTGTTCACCAACAACAG	1378
Qy	1080	ACTGGTCAATACCAACACCTCTTCCCAACA	ACACAGAGAGAGCTCTACACATGACATC	1139
Db	1379	GCTGGTCAATAACAACACCTTTTCCAAACA	ACCTAGTAGAGTCTCTACACATGACATC	1438
Qy	1140	ACCCAGATGGTCAACACCCGAAATCAGATTGA	TTATATTTCTTTGCGAGCCAAAGATGGAAG	1199
Db	1439	ACCAGATGGTTCGATACCGAGATCAGATTGA	TTATATTTTGTGCGCCAAAGATGGAAG	1498
Qy	1200	CTCTATACAGTCAGCAAAAACAAGCACAGG	AGCTTACTGTGGCTCAGATCTATGACTCTCT	1259
Db	1499	CTCTGTACAGTCAGCAAAAACAAGACCTTGA	AGCTGACTGTGGCTCAGATCTATGACTCTCT	1558
Qy	1260	TATTGCCAAATTCAGACTTTAAATGAAGAA	AGTAGGGGAAAACCACTAGATCACTCAG	1316

1559 TATTGCAAGTTTCAGGCTTAAGTTGAAGATATACCAAAAACGACTCGGCATTCAG 1615

Db

RESULT 12

ABSS5699

ID ABS55699 standard; DNA; 54842 BP.

XX

AC ABS55699;

XX

XX 08-JAN-2003 (first entry)

XX

XX Bovine Claudin-16 deficiency associated polynucleotide sequence #2.

XX

XX Bovine; Claudin-16 deficiency; type 2 mutation; ds.

XX

XX Bos sp.

XX

XX JP2002238570-A.

XX

XX 27-AUG-2002.

XX

XX 14-FEB-2001; 2001JP-00037623.

PF

XX

XX 14-FEB-2001; 2001JP-00037623.

XX

XX (CHIK-) CHIKUSAN GIJUTSU KYOKAI SH.

PA

XX (KACH-) KACHIKU KAIRYO JIGYODAN SH.

PA

XX (DOKU-) DOKURITSU GYOSEI HOJIN KACHIKU KAIRYO CE.

XX

XX WPI; 2002-744759/81.

DR

XX

XX

PT

XX

PT Gene diagnosis of type 2 mutation for bovine claudin-16 deficiency,

XX

XX comprises checking for mutation in 1-161 position in claudin-16 sequence.

PS

XX

PS Example 1; Page 9-28; 43pp; Japanese.

XX

XX

CC

XX

CC The invention describes gene diagnosis of type 2 mutation for bovine

CC

CC claudin-16 deficiency. The method involves (a) getting a bovine nucleic

CC

CC acid sample; (b) subjecting the sample to a gene amplification reaction

CC

CC to give a nucleic acid fragment; (c) checking the presence of mutation in

CC

CC the nucleic acid fragment of step (b). The method is used for gene

CC

CC diagnosis of type 2 mutation for bovine claudin-16 deficiency. This

CC

CC sequence represents a polynucleotide from the bovine Claudin-16

CC

CC deficiency associated gene in which the type 2 mutation occurs

XX

SO

XX

SO Sequence 54842 BP; 15478 A; 10714 C; 10200 G; 18336 T; 0 U; 114 Other;

Query Match	17.4%	Score 619.909	DB 1	Length 54842
Best Local Similarity	63.9%	Pred. No. 0.56		
Matches 1166	Conservative 0	Mismatches 561	Indels 99	Gaps 16
QY	1556	TCTGTATCATGGAAGAACGACAGAGCTTCAGAAAAACATCTATTTCTGCTTTTATGACT	1715	
DB	9575	TCTGGATCATAGGAAAGCAAGGAAATCCA -AAGCACATCTATTTCTACTTCATT -ATT	9632	
QY	1716	ATGCAAAAGCCTTTGACTGTGTGGGGTGCACAAATAACTGTGMAAATCTGAAAGGGATGG	1775	
DB	9633	ACATTAAGCCTTTTGACAAATGTGATCAGMAAAATATGGAGATT-----AAAATGG	9689	
QY	1776	GAATACCAGACCACTGACCTGACTCTTGAAAAATTTGTATGCAGGTGAGGAACAACAG	1835	
DB	9686	AAATACCAGACCACTTTACCTGTCTCTGAGAAACCTGTATTTCAGGACAAGAAATCAACAG	9745	
QY	1836	TTAGAACTGGACATGGAAACAACAGACTGGTTTCCAAAGTAGGMAAAGGAGTATGTCAAGGCT	1895	
DB	9746	TTAGAACTGGACATGGAAACAACAGACTGGTTTCAAATGAGGAGAGAAATACGTCAAGGCT	9805	
QY	1896	GTATATTGTTCACCGCGCTTGTTTAACTTCTATGCAGAGACATCATGAGAAACGCTGGCT	1955	
DB	9806	GTATATTGTTCACATTGCTTTATTAACTTTATTTCAGAAATCACTTAT-----	9850	
QY	1956	GGAAAGAACCAAGCTGGAATCAAGATTGCGGGAGAAATAGCAATAAACCCTCAGATATGC	2015	

Db 2693 TTCTGGGAAATATCAATAAAC-GAGATACAAAGAT-ACACCACACTTATGGCAGAAA 2750
 QY 2043 TGAAGAGAACTAAAGACCTCTTGATCAAGGTGAAAGAGAGAGTGAARAAGTTGCTT 2102
 Db 2751 CTAAGAAGAACTAAAGACCTCTTGATGAAGTGAAGAGAGAGTGAARAAGCCAGCTT 2810
 QY 2103 AAAGCTCAACATTCAGAAAAGCAAGATCATGGCATCTGGTCCCATCTCTCATGGAAAT 2162
 Db 2811 AAAACCCAACTCAATCAAGA-----TCATCTTTCATGTGCAAT 2853
 QY 2163 AGATGGGAAACAGTGGAAACAGTGTCTAGCTTTATTTTGGGGGGCTCCAAATCACT 2222
 Db 2854 AAATGGGAAACAAATGGAAACAGTGAAGACTTTATTTCT--TGGGCTCCAAATCACT 2911
 QY 2223 GCAGATGTGATCGAGCCATGAATTAAGACACTTACTCTTGGAGAAAGTTA-- 2280
 Db 2912 CGAGATTGTGACTACAGCCATG--ATTAAAGATGCTTGCTCTTGAAGAGAGCTATT 2969
 QY 2281 ACCAACCCTAGATAGATATTGAAAGCAGAGACATTACCTTGGCCAAACAAAGCCCATCTA 2340
 Db 2970 ACCAAACTAGAAAGCATATTAAAGCAGAGACGTTACTTGTCTGACTAAAGTTCTGTCTA 3029
 QY 2341 GTCAGGCTATGTTTTTCCAGTGGTCTATGTATGATGTGAGAGTGTGACTGTGAAGAA 2400
 Db 3030 GTCAAACCTATGGTTTTTCCAGTGTCTATATATGATGT--GAGTTGAACCTATAAGAAA 3087
 QY 2401 GCTGAGCACTGAAGAATTGATGCTTTTGAACCTGTGGTGTGGAGAGACTCTTGAGAGTC 2460
 Db 3088 GCTGAGCACCAAGAAATTGATGCTTTTGAATTTGGTGTGGAGAGTCTCTTGAGAGTC 3147
 QY 2461 CTTTGA--CTGCAAGAGATCAACACAGTCAATCTGAAGAGATCAGCCCTGGATTTC 2519
 Db 3148 CTTTGAACCTCAAGAGATCCAAACAGTCCATCTTAAGGAAATCAGTCTTGAATATTC 3207
 QY 2520 TTTTGAAGCAATGATGCTTAAAGCTGAAACTCCAGTACTTTGGCCACCTGA--TCAGAA 2577
 Db 3208 ATTGGAGACTGATGCTGAAATTGAGATTAAGTTTGGACTCACTTAATGCAGAGA 3267
 QY 2578 GCTGACTCACTGGAAGACCTGATGCTGGAGGGATTTGGGGCAGAGAGAGAGGGA 2637
 Db 3268 GCCAACTCACTAGAAAAGACCC-CATGTTGGCAAAAATTGAAGCCAGAGAGAGAGTGA 3326
 QY 2638 CGACAGAGATGAGATGCTGGATGGCATCACTGATGAGCGTGAAGTCTGGGTGA 2697
 Db 3327 TGACAGAGATGAGATGCTTGGATGGCATCGTTGACTGAATGGACATGATCTGATCAAG 3386
 QY 2698 CTCCTGGAGTTGGTATGACAGAGAGCCCTGCTCTCGCGGATTCATGGGGTCAAAAG 2757
 Db 3387 TTCGGGAGACAGCAAGGACAGGGCTGCTGTGCTGTCAGTCCATGGGGTGCAGAG 3446
 QY 2758 AGTTGGACCACTGAGCAACTGAAC 2783
 Db 3447 AGTCGGTCTCAATGAGTAACATAAAC 3472

RESULT 15
 AAF77439

ID AAF77439 standard; DNA; 13723 BP.

XX
 AC AAF77439;

DT 15-JUN-2001 (first entry)

DE Goat beta-casein gene.
 XX

KW Goat; beta-casein; ds.
 XX

OS Capra hircus.
 XX

FN TW413702-A.
 XX

PD 01-DEC-2000.
 XX

PF 28-MAY-1997; 97TW-00107261.
 XX
 PR 28-MAY-1997; 97TW-00107261.
 XX
 PA (NASC-) NAT SCI COUNCIL.
 XX
 PI Huang M, Lin J;
 XX
 DR WPI; 2001-264857/27.
 XX
 PT DNA molecules containing goat beta-casein gene --
 XX
 PS Claim 1; Page 1-13; 41pp; Chinese.
 XX
 CC This invention relates to the goat beta-casein gene, represented by the
 CC present sequence. There are 18 positions in the 5'-flanking region, 21
 CC positions in intron sequences and 1 position in exon VII that conflict
 CC with the sequence reported for this gene by Roberts et al (1992)
 XX
 SQ Sequence 13723 BP; 4471 A; 2396 C; 2470 G; 4386 T; 0 U; 0 Other;
 Query Match 13.8%; Score 494.6; DB 1; Length 13723;
 Best Local Similarity 90.6%; Pred. No. 5.6;
 Matches 538; Conservative 0; Mismatches 54; Indels 2; Gaps 1;
 QY 2220 ACTGAGATGTTGATCTGAGCCATGAAATTAAGACACTTACTCTTGGAGAAAGTT 2279
 Db 8768 ACTAAGAATGATTATTGAGCCATGAAATTAAGACACTTACTCTTGGAGAAAGTT 8827
 QY 2280 A--ACCACTAGATAGATATTGAAGCAGAGACATTACTCTCCCAACAAAGCCCAT 2337
 Db 8828 ATGACCACTAGATAGATATTGAAGCAGAGACATTACTCTCCCACTAAGTCCGT 8887
 QY 2338 CTAGCAAGCTATGTTTTCAGTGTCTATGTATGATGTGAGATTGACCTGTGAAG 2397
 Db 8888 CTAGTCAGGCTATGTTTTCAGTGTCTATGTATGATGTGAGATTGACCTGTGAAG 8947
 QY 2398 AAAGCTGAGCACTGAAGAATTGATGCTTTTGAACCTGTGTTGGAGAGACTCTTGAGA 2457
 Db 8948 AAGGCTGAGTGCCAAAGAAATAGATGCTTTTGAACCTGTGTTGGAGAGACTCTTGAGA 9007
 QY 2458 GTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCTGAAAGGAGATCAGCCCTGGGATT 2517
 Db 9008 GTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCTGAAAGGAGATCAGCCCTGGGATT 9067
 QY 2518 TCTTTGGAAGGAATGATGCTAAAGCTGAACCTCCAGTACTTTGGCCACCTCTTCGAGA 2577
 Db 9068 TCTTTGGAAGGAATGATGCTAAAGCTGAACCTCCAGTACTTTGGCCACCTCTTCGAGA 9127
 QY 2578 GCTGACTCACTGAAAAGACCTGATGCTGGAGGGATTGGGGCAGAGAGAGAGGGA 2637
 Db 9128 GTTGACTCAATTGAAAAGACTCTGATGCTGGAGGGATTGGGGTCAAGAGAGAGGGA 9187
 QY 2638 CGACAGAGATGAGATGCTGGATGGCATCACTGATCGATGAGCTGAGTCTGGGTGAA 2697
 Db 9188 CGACCGAGGATGAGATGCTGGATGGCATCACTGATCGATGAGCTGAGTCTGGGTGAA 9247
 QY 2698 CTCCTGGAGTTGGTATGAGCAGAGGGAGCCCTGCTCTCGGGGATTCATGGGGTCAAG 2757
 Db 9248 CTCGGGAGTTGGTATGAGCAGAGGGAGCCCTGGTGTCTGCAATTCATGGGGTTCGAG 9307
 QY 2758 AGTTGGACCACTGAGCAACTGAACCTGAACCTGACTGACTGAACCTGATTAGTA 2811
 Db 9308 AGTCAGACACGACTGAGTGAAGTGAACCTGAACCTGAACCTGAAGAAATGATTA 9361

RESULT 16
 ABN74558/c

ID ABN74558 standard; cDNA; 908 BP.

XX
 AC ABN74558;

XX
 DT 03-JUL-2002 (first entry)

Db 1056 CCGTCTAGTCAAGGCTATGCTTTTCCAGTAGTCATGTATGGAGTGCAGAGTTGGACTGT 997
 Qy 2394 GAAGAAAGCTGAGCACTGAAGAAATGATGCTTTTGAAGTGGTGTTGGAGAAAGACTCTT 2453
 Db 996 GAAGAAAGCTGAGGCTGAAGAAATGATGCTTTTGAAGTGGTGTTGGAGAAAGACTCTT 937
 Qy 2454 GAGAGTCCCTTGGAGTGAAGAGATCCAAACAGTCCATTCTGAGAGATCAGCCCTGG 2513
 Db 936 GAGAGTCCCTTGGAGTGAAGAGATCCAAACAGTCCATTCTGAGAGATCAGCCCTGG 877
 Qy 2514 CATTTCTTTTGAAGGAATGATGCTTAAAGCTGAAACTCCAGTACTTTTGGCCACCTGATCAG 2573
 Db 876 GATTCTTTTGAAGGAATGATGCTTAAAGCTGAAACTCCAGTACTTTTGGCCACCTCAGCA 817
 Qy 2574 AAGAGTCACTACCTGGAAGAACCCCTGATGCTGGAGGGATTTGGGGGCGAGGAGAAG 2633
 Db 816 AAGAGTTGACTCATTTGGAAAGACTCTGATGCTGGAGGAACCTGGGGGCGAGGAGAAG 757
 Qy 2634 GGGACGACAGAGGATGAGTGGCTGGATGGATCACTGATCGATGACGTGAGTCTGGG 2693
 Db 756 GGGACACAGAGGATGAGTGGCTGGACGGATCCTACTCGATGACGTGAGTCTGAG 697
 Qy 2694 TGAACCTCTGAGTGGTGTGATGGACAGGAGGCTGCTCTGCGCGGATTCATGGGFTAC 2753
 Db 696 TGAACCTCTGGAGTTGGTGTGATGGACAGGAGGCTGCTCGATTCGATTCATGGGFTCG 637
 Qy 2754 AAGAGTTGGACAGACTGACCACTGAACCTGAACCTGAACCTGAACCTGAACCTTAGT 2813
 Db 636 AAGAGTTCAGACAGACTGACGCACTGAACCTGAACCTGAACCTGAACCTGAACCTGTTACAGA 577
 Qy 2814 TTATATTACTCAGAAAATAGTAATTTTCATATGTATTTCAAAATATTTTCATA 2864
 Db 576 AATTACAGGAAAATAATGATCTTTGTTAAACATTTGCAATTTTGTAAA 526

RESULT 18

AAD26492/c
 ID AAD26492 standard; DNA; 1813 BP.
 AC AAD26492;
 DT 26-MAR-2002 (first entry)
 DE Sheep alpha (1, 3) galactosyltransferase post-exon 4 DNA #3.
 KW Sheep; xenotransplantation; Gal-alpha (1,3)Gal; GAL determinant; surgery;
 KW alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;
 KW gene therapy; ds.
 XX Ovis sp.
 OS
 XX WO200188096-A2.
 FN
 XX 22-NOV-2001.
 PD
 XX 14-MAY-2001; 2001WO-US015765.
 PF
 XX 15-MAY-2000; 2000US-0204148P.
 PR
 XX 13-JUN-2000; 2000US-00593316.
 XX (GERO-) GERON CORP.
 PA
 XX Denning C, Clark J;
 FI
 XX WPI; 2002-089848/12.
 DR
 XX New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal
 PT determinants for xenotransplantation, and in the treatment of the human
 PT body by surgery or therapy.
 XX
 XX Claim 18; Page 79; 86pp; English.
 PS
 XX The patent discloses immunologically compatible animal tissue, suitable

CC for xenotransplantation into human patients. The invention also relates
 CC to an ovine tissue devoid of antibody-detectable Gal-alpha (1,3)Gal (GAL)
 CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The
 CC ovine tissue is useful for treatment of human body by surgery or therapy
 CC and in xenotransplantation, by transplanting the ovine tissue into a
 CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.
 CC Polynucleotide constructs of the invention are useful for inactivating an
 CC alpha(1,3) GT gene in an ovine cell. Alpha (1,3) GT sequences are useful
 CC for drug screening and for the production of GAL containing synthetic
 CC oligosaccharides. Sequences of the invention are also useful in gene
 CC therapy. The present sequence is alpha (1,3) GT post-exon 4 DNA
 XX
 SQ Sequence 1813 BP; 477 A; 391 C; 428 G; 515 T; 0 U; 2 Other;
 Query Match 10.9%; Score 388.9; DB 1; Length 1813;
 Best Local Similarity 88.2%; Pred. No. 88;
 Matches 455; Conservative 0; Mismatches 56; Indels 5; Gaps 3;
 Qy 2282 CCAACTAGATAGCATATTGAAAGCAGAGACATTACCTTGCACAAAGCCCATCTA- 2340
 Db 936 CTAAATGTTTATAGCATATTCAAAACAGAGACATTACTTTGCCACAAAGGTCCTAT 877
 Qy 2341 GTCAAGGCTATGTTTTCAGTGTCTATGTATGGATGTGAGTGTGACTGTGAAGAA 2400
 Db 876 GTCAAGGCTATGTTTTCAGTGTCTATGTATGGATGTGAGTGTGACTGTGAAGAA 817
 Qy 2401 GCTGAGCACTGAAGATTTGCTTTTGAAGTGTGTTTGGAGNAGACTCTTGAGAGTC 2460
 Db 816 GCTGAGCCCCGAGAAATCGATGCTTTTGAAGTGTGTTTGGAGNAGACTCTTGAGAGTC 757
 Qy 2461 CTTTGGAGCTGCAAGGAGATCCAAACAGTCCATTCTTGAAGAGAGATCAGCCCTGGATTCT 2520
 Db 756 CTTTGGAGCTGCAAGGAGATCCAAACAGTCCATTCTTGAAGAGAGATCAGCCCTGGATTCT 697
 Qy 2521 TTGGAAGAGATGATCTAAAGCTGAACTCCAGTACTTTGGCCACCTGATCAGAGAGCT 2580
 Db 696 TTGGAAGAGATGATCTAAAGCTGAACTCCAGTACTTAAGGCCACCTCATGAGAGAGT 637
 Qy 2581 ---GACTCACTGGAAGAACCTGATGCTGGAGGGATTTGGGGGCGAGGAGAGAGGGA 2637
 Db 636 TGACTCATATGGAAGAGACTCATGAAGCTGGAGGGATTTGGGGGCGAGGAGAGAGGGA 577
 Qy 2638 CGACAGAGATGATGCTGGATGGCATCTGATCGATGAGTGCAGTGCAGTCTGGTGAA 2697
 Db 576 CGACAGAGATGATGCTGGATGGCATCTGATCGATGAGTGCAGTGCAGTCTGGTGAA 517
 Qy 2698 CTCTCGGAGTTGGTGTGATGACAGAGGAGCCCTGCTCGGGCGATT-CATGGGGTCACAAA 2756
 Db 516 CTCTAGGAGTTGGTGTGATGACAGAGGAGCCCTGGCGTCTGCATTCATGGGGTCGAAA 457
 Qy 2757 GAGTTGGACAGCACTGACCAACTGAACCTGAACCTGAA 2792
 Db 456 GAGTGGACAGCACTGAACGAGCTGAACCTGAACCTGAA 421

RESULT 19

ADE81111
 ID ADE81111 standard; DNA; 967 BP.
 XX
 AC ADE81111;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Bovine 1-8U protein-encoding gene upstream region.
 XX
 KW bovine 1-8U protein; bovine Leu-13 protein; interferon; bovine pregnancy;
 KW detecting pregnancy; ds.
 XX
 XX Bos taurus.
 OS
 XX US2003143601-A1.
 FN
 XX 31-JUL-2003.


```
XX 18-NOV-2002; 2002US-00299497.
XX 16-NOV-2001; 2001US-0337090P.
XX (HANS/) HANSEN T R.
XX (AUST/) AUSTIN K J.
XX (PRUO/) PRU J K.
XX Hansen TR, Austin KJ, Pru JK;
XX WPI; 2003-829789/77.
XX Novel isolated bovine 1-8U protein and Leu-13 protein whose expression is
XX induced by interferon and elevated during bovine pregnancy, useful for
XX detecting pregnancy in bovine animal.
XX Claim 13; SEQ ID NO 11; 33pp; English.
XX This invention relates to novel isolated bovine 1-8U and Leu-13 proteins
XX whose expression is induced by interferon and elevated during bovine
XX pregnancy. The invention is useful for detecting pregnancy in a bovine
XX animal. The method involves taking biological samples from a test animal
XX and from a non-pregnant animal and detecting bovine 1-8U protein or Leu-
XX 13 protein in which an elevation of the protein from the test animal
XX relative to that from non-pregnant animal indicates pregnancy in the test
XX animal. The present sequence is that of the upstream region of the gene
XX which encodes the bovine 1-8U protein and which is related to the
XX invention.
XX Sequence 967 BP; 313 A; 173 C; 238 G; 243 T; 0 U; 0 Other;
XX
XX Query Match 10.78; Score 380.6; DB 1; Length 967;
XX Best Local Similarity 81.08; Pred. No. 1.7e+02;
XX Matches 443; Conservative 0; Mismatches 104; Indels 0; Gaps 0;
XX
QY 2251 AAAGACACTTACTCTCTGGAAGAAAGTTAAACCAACCTAGATAGCATATTGAAACGAGA 2310
DB 297 AAAGGGAATTGCAAAAATATAAAAATGTAATAAAATGTAGACAGCATATTAAAAGCAGA 356
QY 2311 GACATTACCTTGCACAAAGCCCATCTAGTCAGGCTATGTTTCCAGTGTCATG 2370
DB 357 GACATTACTTGTGCGTAAAGGTCATATAGTCAAGCTATGTTTCCAGTAGTCAIG 416
QY 2371 TATGATGTGAGAGTGTGGAAGTGTGGAAGAAAGCTGAGCACTGAAAGATTTGATGTTTGA 2430
DB 417 GATGATGTGAGAGTGTGGAAGTGTGGAAGAAAGCTGAAATGCCGAGAAATTTGATGTTGA 476
QY 2431 CTGTGTTGTTGGAAGAGACTCTTGAGAGTCCCTTTGGAAGTCCGAGGAGATCCACCAAGTCC 2490
DB 477 CTGTGTTGTTGGAAGAGACTCTTGAGAGTCCCTTTGGAAGTCCGAGGAGATCCACCAAGTCA 536
QY 2491 ATTCGAGGAGATCAGCCCTGGGATTTCTTTGGAAGGATGATGCTAAAGCTGAAACTC 2550
DB 537 ATCTTAAATTAATCAGCCCTTAATTTGATTTGGAAGGATGTTGCTGAAACTGAAACTC 596
QY 2551 CAGTACTTTGGCCACCTGATCAGAGAGCTGACTCCTGGAAGAACCTCTGATGCTGGGA 2610
DB 597 CAATACCTTTGGCTACCTGATGATCAAGAACTGACTCATTGGAAGAACCTCTGATGCTAGGA 656
QY 2611 GGGATTGGGCGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2670
DB 657 AAGATTGAAGCGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 716
QY 2671 GACTCGATGAGGAGTGTGAGTCTGGGTGAATCTCTGAGTTGGTGTGAGGAGGAGGAGGAGGAG 2730
DB 717 GACTCGATGAGGAGTGTGAGTCTGGGTGAATCTCTGAGTTGGTGTGAGGAGGAGGAGGAGGAG 776
QY 2731 CCTCGGCGATTCATGGGGTCAAAAGAGTTGGGACAGCATGAGCAACTGAATGAAGTCTG 2790
DB 777 AGTGTGCGAGTCCACAGGGTGCAGAGAGTGCAGACTACTGAGCGAGTGAAGTGAAGTGA 836
QY 2791 AACTGTA 2797
```

```
Db 837 AACTGAA 843
RESULT 20
ABX48667
ID ABX48667 standard; cDNA; 419 BP.
XX AC ABX48667;
XX DT 21-FEB-2003 (first entry)
XX DE Bovine EST associated with lactation/muscle/fat deposition #13832.
XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
XX KW muscle deposition; fat deposition; genome mapping; gene identification;
XX OS Bos Taurus.
XX PN US2002137139-A1.
XX PD 26-SEP-2002.
XX PF 24-SEP-2001; 2001US-00960352.
XX PR 12-JAN-1999; 99US-0115707P.
XX PR 11-JAN-2000; 2000US-00480902.
XX PA (BYAT/) BYATT J C.
XX PA (MATH/) MATHIALAGAN N.
XX PA (TAON/) TAO N.
XX PA (WARR/) WARREN W C.
XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WIPI; 2003-110599/10.
XX New nucleic acid associated with lactation, and muscle and fat
XX deposition, useful for genome mapping, gene identification and analysis,
XX cattle breeding, or for genetically improving cattle.
XX Claim 2; SEQ ID NO 13832; 245pp; English.
XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMFD), derived from
XX cattle, and the LMFD nucleic acid can specifically hybridize to a second
XX nucleic acid molecule comprising any of 15112 nucleotide sequences,
XX appearing as ABX34836-ABX49947, or complements of them. Also included are
XX ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
XX acid linked to a promoter and a 3' non-translated sequence that
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX (2) determining a level or pattern of a molecule in a bovine cell or
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX of the 15112 nucleic acid sequences or its complement or fragment) with a
XX complementary nucleic acid molecule obtained from the bovine cell or
XX tissue, where hybridisation between the marker nucleic acid and the
XX complementary nucleic acid permits the detection of the molecule; and (b)
XX detecting the level or pattern of the complementary nucleic acid, where
XX the detection of the complementary nucleic acid is predictive of the
XX level or pattern of the molecule. The LMFD nucleic acid is used for
XX determining a level or pattern of a molecule in a bovine cell or tissue.
XX It is useful for genome mapping, gene identification and analysis, cattle
XX breeding, preparation of constructs for use in cattle gene expression, or
XX for genetically improving cattle. The present sequence is one of the
XX 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
XX present sequence was not shown in the specification but was obtained in
XX electronic format from the USPTO web site:
XX seqdata.uspto.gov/sequence.html?DocID=20020137139
XX Sequence 419 BP; 105 A; 71 C; 143 G; 100 T; 0 U; 0 Other;
```


XX Bovine EST associated with lactation/muscle/fat deposition #6161.
DE Bovine, ss; EST; expressed sequence tag; lactation; LMFD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX Bos Taurus.
XX US2002137139-A1.
XX 26-SEP-2002.
XX 24-SEP-2001; 2001US-00960352.
XX 12-JAN-1999; 99US-0115707P.
XX 11-JAN-2000; 2000US-00480902.
XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARR/) WARREN W C.
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX New nucleic acid associated with lactation, and muscle and fat
XX deposition, useful for genome mapping, gene identification and analysis,
XX cattle breeding, or for genetically improving cattle.
XX Claim 2; SEQ ID NO 6161; 245pp; English.
XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMFD), derived from
XX cattle, and the LMFD nucleic acid can specifically hybridize to a second
XX nucleic acid molecule comprising any of 15112 nucleotide sequences,
XX appearing as ABX34836-ABX49947, or complements of them. Also included are
XX ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
XX acid linked to a promoter and a 3' non-translated sequence that
XX functions in the cell to cause termination of transcription and addition
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX (2) determining a level or pattern of a molecule in a bovine cell or
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX of the 15112 nucleic acid sequences or its complement or fragment) with a
XX complementary nucleic acid molecule obtained from the bovine cell or
XX tissue, where hybridisation between the marker nucleic acid and the
XX complementary nucleic acid permits the detection of the molecule; and (b)
XX detecting the level or pattern of the complementary nucleic acid, where
XX the detection of the complementary nucleic acid is predictive of the
XX level or pattern of the molecule. The LMFD nucleic acid is used for
XX determining a level or pattern of a molecule in a bovine cell or tissue.
XX It is useful for genome mapping, gene identification and analysis, cattle
XX breeding, preparation of constructs for use in cattle gene expression, or
XX for genetically improving cattle. The present sequence is one of the
XX 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
XX present sequence was not shown in the specification but was obtained in
XX electronic format from the USFTO web site:
XX seqdata.uspto.gov/sequence.html?DocID=20020137139

XX Sequence 436 BP; 156 A; 84 C; 96 G; 100 T; 0 U; 0 Other;
Query Match 10.2%; Score 363.2; DB 1; Length 436;
Best Local Similarity 93.4%; Pred. No. 4.3e+02;
Matches 410; Conservative 0; Mismatches 23; Indels 6; Gaps 3;
XX 1992 GGCCTGTATTGTCCACCGGCTGCTTTAACTTCTATGCAGAG-ACATCATGAGAAAGCCT 1950
DB 1 GGCAGTATTGTTCACCTGCTTTATTTAACTTCTATGCAGAGTACATCATGAGAAAGCCT 60
XX 1951 GGGCTGGAGAGACACAGCTGGATCAGATTCCCGGAGAAATAGCATACCTCAGA 2010
DB 61 GGACTGGAGAAACACAGCTGGAAATCAAGATTCCCGGAGAAATATCAATAACCTCAGA 120

QY 2011 TATGCAGATGATACACCCCTTATGGCAGAAAAGTGAAGAGGAACCTAAAGAGCCTCTTGATG 2070
DB 121 TATGCAGATGATACACCCCTTATAGCAGAAAAGTGAAGAGGAACCTAAAGAGCCTCTTGATG 180
QY 2071 AAGGTGAAAGAGAGAGAGTGAAGAAAAGTGGCTTTAAAGCTCAACATTCAGAAAACCAAGATC 2130
DB 181 AAAGTGAATTTGGAGAGTGAAGAAAAGTGGCTTTAAAGTTCACATTCAGAAAATGAAGATC 240
QY 2131 ATGGCATCTGTCCTCCATCACCTTCATGGGAAATAGATGGGAAAACAGTGGAAACAGTGTCA 2190
DB 241 ATGGCATCTGTCCTCCATCACCTTCATGGGAAATAGATGGGAAAACAGTGGAAACAGTGTCA 300
QY 2191 CACTTTATTTTGGGGGGCTCCAAAATCACTGCAGATGGTGAAGTGGAGCCATGAATTA 2250
DB 301 CACTTTATTTT---GGGCTCCAAAATCACTGCAGATGGTGAAGTGGAGCCATGAATTA 357
QY 2251 AAAGACACTTACTCTCTGGGAGAAAAGTTA--ACCACTAGATAGATATTGAAAAGCA 2308
DB 358 AAAGACGCTTACTCTCTGGGAGAAAAGTTATGACCACTAGATAGATATTGAAAAGCA 417
QY 2309 GAGACATTACCTTGCCAAAC 2327
DB 418 GAGACATTACTTTGGCAAAC 436
RESULT 23
ABX39925
ID ABX39925 standard; cDNA; 432 BP.
XX AC ABX39925;
XX DT 20-FEB-2003 (first entry)
XX DE Bovine EST associated with lactation/muscle/fat deposition #5090.
XX KW Bovine, ss; EST; expressed sequence tag; lactation; LMFD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
XX Gene analysis; cattle breeding.
XX OS Bos Taurus.
XX PN US2002137139-A1.
XX PD 26-SEP-2002.
XX PF 24-SEP-2001; 2001US-00960352.
XX PR 12-JAN-1999; 99US-0115707P.
XX PR 11-JAN-2000; 2000US-00480902.
XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARR/) WARREN W C.
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX New nucleic acid associated with lactation, and muscle and fat
XX deposition, useful for genome mapping, gene identification and analysis,
XX cattle breeding, or for genetically improving cattle.
XX Claim 2; SEQ ID NO 5090; 245pp; English.
XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMFD), derived from
XX cattle, and the LMFD nucleic acid can specifically hybridize to a second
XX nucleic acid molecule comprising any of 15112 nucleotide sequences,
XX appearing as ABX34836-ABX49947, or complements of them. Also included are
XX ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
XX acid linked to a promoter and a 3' non-translated sequence that
XX functions in the cell to cause termination of transcription and addition
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX (2) determining a level or pattern of a molecule in a bovine cell or
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX of the 15112 nucleic acid sequences or its complement or fragment) with a
XX complementary nucleic acid molecule obtained from the bovine cell or
XX tissue, where hybridisation between the marker nucleic acid and the
XX complementary nucleic acid permits the detection of the molecule; and (b)
XX detecting the level or pattern of the complementary nucleic acid, where
XX the detection of the complementary nucleic acid is predictive of the
XX level or pattern of the molecule. The LMFD nucleic acid is used for
XX determining a level or pattern of a molecule in a bovine cell or tissue.
XX It is useful for genome mapping, gene identification and analysis, cattle
XX breeding, preparation of constructs for use in cattle gene expression, or
XX for genetically improving cattle. The present sequence is one of the
XX 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
XX present sequence was not shown in the specification but was obtained in
XX electronic format from the USFTO web site:
XX seqdata.uspto.gov/sequence.html?DocID=20020137139

CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 CC Sequence 432 BP; 146 A; 87 C; 100 G; 99 T; 0 U; 0 Other;

Query Match 10.1%; Score 360.3; DB 1; Length 432;
 Best Local Similarity 94.3%; Pred. No. 4.4e+02;
 Matches 384; Conservative 0; Mismatches 22; Indels 1; Gaps 1;
 Qy 1589 GAACCTCCAGTTGTTCAAGCTGGTTTGTAGAAAAGTCAGAGAACCCAGACCAATGGCC 1648
 Db 26 GATCTTCCAGTTGTTCAAGCTGGTTTGTAGAAAAGTCAGAGAACCCAGACCAATGGCC 85
 Qy 1649 AACATCTCTGTATCATGGAAGAGAGAGTTCCAGAAAACATCTATTTCTGCTTT 1708
 Db 86 AACATCTCTGTATCATGGAAGAGAGAGTTCCAGAAAACATCTATTTCTGCTTT 145
 Qy 1709 ATTGACTATGCAAAAGCCTTTGACTGTGGGGTCCACAAATAACTGTGGAATTTCTGAAA 1768
 Db 146 ATTGACTATGCAAAAGCCTTTGACTGTGGGGTCCACAAATAACTGTGGAATTTCTGAAA 205
 Qy 1769 GGGATGGGAATACAGACCACTGACTCTGTTGAAAAATTTGTATGAGTCCAGGAA 1828
 Db 206 AAGATGGGAATACAGACCACTGACTCTGTTGAAAAATTTGTATGAGTCCAGGAA 265
 Qy 1829 GCACAGTTAGAACTGGACATGGACACACAGACTGGTTCCAGTACGAAAGAGTATCT 1888
 Db 266 GCACAGTTAGAACTGGACATGGACACACAGACTGGTTCCAGTACGAAAGAGTATCT 325
 Qy 1889 CAAGGCTGTATATGTGTCACCGGCTTTGTTAACTTCTATGTCAGAG-ACATCATGAGAAC 1947
 Db 326 CAAGGCTGTATATGTGTCACCGGCTTTGTTAACTTCTATGTCAGAGTACATTAAGAAC 385
 Qy 1948 CTGGGCTGGAAGAGACAGACTGGAAATCAAGATTGCCGGGAGAA 1994
 Db 386 CTGGGCTGGAAGAGACAGACTGGAAATCAAGATTGCCGGGAGAA 432

RESULT 24
 ID ABX39519
 XX ABX39519 standard; cDNA; 433 BP.
 XX
 AC ABX39519;
 XX
 XX
 DT 20-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #4684.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 PN US2002137139-A1.
 XX

PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-00960352.
 XX
 PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX
 (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX
 Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX
 WPI; 2003-110599/10.
 DR
 XX
 PT New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX
 PS Claim 2; SEQ ID NO 4684; 245pp; English.
 XX
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC: (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 SQ Sequence 433 BP; 153 A; 74 C; 105 G; 101 T; 0 U; 0 Other;
 Query Match 10.0%; Score 357.1; DB 1; Length 433;
 Best Local Similarity 90.5%; Pred. No. 4.5e+02;
 Matches 401; Conservative 0; Mismatches 29; Indels 13; Gaps 2;
 Qy 1943 GAACGCTGGCTGGAGAGACCAAGCTGGAATCAGATTCGCGGAGAAATAGCAATA 2002
 Db 1 GAACGCTGGCTGGAGAGACCAAGCTGGAATCAGATTCGCGGAGAAATAGCAATA 60
 Qy 2003 ACCTCAGATATGCAGATGATACCAACCTTATGGCAGAAAGTGAAGAGAGAACTAAAAAGCC 2062
 Db 61 ACCTCAGATATGCAGATGATACCACTACCTTTATGGCAGAAAGTGAAGAGAGAACTAAAAAGCC 120
 Qy 2063 TCTTGATCAGAGTCAAGAGAGAGAGTCAAAAAGTTGGCTTAAAGCTCAACATTCAGAAA 2122
 Db 121 TCTTGATCAGAGTCAAGAGAGAGAGTCAAAAAGTTGGCTTAAAGCTCAACATTCAGAAA 180
 Qy 2123 CGAAGATCATGGCATCTGGTCCCATCTCTCATGGGAAATAGATGGGAAACAGTGGAA 2182
 Db 181 CTAAGATCATGGCATCTGGTCCCATCTCTCATGGGAAATAGATGGGAAACAGTGG- 237
 Qy 2183 CAGTGTGAGACTTTATTTTGGGGGCTCCAAATCAGTGCAGATGCTGACTGCAGCCA 2242

Db 238 -----CTGACTTTTATTTGGGGGCTCCAAAATCACTGCAGATGGTAACCTGAGGCCA 289
 QY 2243 TGAATTAAGACACCTTACTCTTGGAAAGAAAGTTA--ACCAACCTAGATAGCATATT 2300
 Db 290 TGAATTAAGACACCTTACTCTTGGAAAGAAAGTTA--ACCAACCTAGATAGCATATT 349
 QY 2301 GAAAGCAGACATTAACCTTGGCAACAAAGCCCATCTAGTCAAGGCTATGTTTTTCC 2360
 Db 350 AATAGCAGACATCACTTTGTCAACAAAGGTCATCTAGTCAAGGCTATGTTTTTTC 409
 QY 2361 AGTGGTCATGATGGATGTGAGA 2383
 Db 410 AGTGGTCATGATGATGTGAGA 432

RESULT 25

ABX40182
 ID ABX40182 standard; cDNA; 433 BP.

XX AC ABX40182;

XX DT 20-FEB-2003 (first entry)

XX DE Bovine EST associated with lactation/muscle/fat deposition #5347.

XX XW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 XW muscle deposition; fat deposition; genome mapping; gene identification;
 XW gene analysis; cattle breeding.

XX OS Bos Taurus.

XX PN US2002137139-A1.

XX PD 26-SEP-2002.

XX XX 24-SEP-2001; 2001US-00960352.

XX XX 12-JAN-1999; 99US-0115707P.

XX PR 11-JAN-2000; 2000US-00480902.

XX XX (BYAT/) BYATT J C.

XX PA (MATH/) MATHIALAGAN N.

XX PA (TAON/) TAO N.

XX PA (WARR/) WARREN W C.

XX FI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX DR WPI; 2003-110599/10.

XX PT New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.

XX PS Claim 2; SEQ ID NO 5347; 245pp; English.

XX CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for

CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139

XX SQ Sequence 433 BP; 159 A; 81 C; 94 G; 99 T; 0 U; 0 Other;

Query Match 9.9%; Score 353.3; DB 1; Length 433;
 Best Local Similarity 90.0%; Pred. No. 4.6e+02;
 Matches 389; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

QY 1701 TCTGCTTTATGACTATGCATAAAGCCCTTGACTCTGGGGTCCACAATAAAGCTGGAAAA 1760
 Db 2 TCTGCTTTATGACTATGCATAAAGCTTTGACTGTGGATCAACAATAAAGCTGGAAAA 61
 QY 1761 TTCTGAAAGGATGGGAATACCAAGACACCTGACTCTTGAAAAATTTGTATGCAG 1820
 Db 62 TTCTGAAAGATGGGAATACCTGCCACCTGACTGCCTCTTGAGAAACCTATATGCAG 121
 QY 1821 GTCAGGAAGCAACAGTTAGAACTGGACATGGAACACACAGACTGGTTCAGTAGAAAAAG 1880
 Db 122 GTCAGGAAGCAACAGTTAGAACTGGACATGGAACACACAGACTGGTTCAGAAAAAG 181
 QY 1881 GAGTATGTCAGAGGCTGTATATTGTACCCCGCTTTTAACTTCTATGCAGAG-ACATCA 1939
 Db 182 GAGTACGTCAGAGGCTGTATATTGTACCCCGCTTTTAACTTCTATGCAGAGTACATCA 241
 QY 1940 TGAGAAAGCGCTGGGCTGGGAAGCAACAGCTGGGAATCAAGATGCCGGGAGAAATAGCA 1999
 Db 242 TGAGAAATGCTGGACTGGATGAAGCCCAAGCTGGAGTCAAAATTCGCGGGAGAGATATCA 301
 QY 2000 ATAACTCAGATATGCAGATATACCACTTATGGCAGAAAGTGAAGGAACTAAAAA 2059
 Db 302 ATAACTCAGATATGCAGATATACCACTTATGGCAGAAAGTGAAGGAAATAAAAA 361
 QY 2060 GCCTCTTGATGAAGCTGAAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGA 2119
 Db 362 GCCTCTTGATGAAGTAAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTTCACATTCAGA 421
 QY 2120 AAACCAAGATCA 2131
 Db 422 AAACCAAGATCA 433

RESULT 26

AAH21025/c

ID AAH21025 standard; DNA; 1848 BP.

XX AC AAH21025;

XX DT 31-AUG-2001 (first entry)

XX DE Bovine-derived DNA fragment contig1 445_67.

XX KW Bovine; digital DNA signature; breeding; animal product origin;
 KW identification; genetic association; breed; population; race; ds.

XX OS Bos taurus.

XX PN DE19959751-A1.

XX PD 13-JUN-2001.

XX XX 11-DEC-1999; 99DE-01059751.

XX PR 11-DEC-1999; 99DE-01059751.

XX XX (FRIE/) FRIES H R.

XX PA (DURS/) DURSTENITZ G.

XX Pries HR, Durstewitz G;
 XX WPI; 2001-376309/40.
 XX New bovine genomic DNA sequences, useful for establishing generic
 PT signatures, e.g. for breeding control, contain specific variable
 PT positions.
 XX Claim 1; Page 22-23; 26pp; German.
 XX This invention describes bovine DNA sequences (A) which are used in a
 CC method to establish a digital, standardized DNA signature. DNA signatures
 CC established from (A) are used to monitor breeding; to determine origin of
 CC animal products; to identify individual animals; to study genetic
 CC association and to establish signatures that are specific at the level of
 CC breed, population or race. The method is based on individual base
 CC exchanges in DNA, and these are inherited more stably (by an order of
 CC magnitude) than conventional markers. It is suitable for population-wide
 CC studies (as a high throughput test) and the presence of unequivocal
 CC sequences flanking the variable position provides a built-in
 CC standardization feature
 XX Sequence 1848 BP; 513 A; 413 C; 407 G; 467 T; 0 U; 48 Other;
 XX
 XX Query Match 9.8%; Score 348.7; DB 1; Length 1848;
 XX Best Local Similarity 86.4%; Pred. No. 1.1e+02;
 XX Matches 418; Conservative 0; Mismatches 63; Indels 3; Gaps 3;
 QY 2312 ACATTACTTCCCAACAAAGCCCATCTAGTCAAGGTATGTTTCCAGTGGTCATGT 2371
 Db 1198 ACATTACTTCCCAACAAAGTCCGCTAGTCAAGGTATGTTTCCAGTGGTCATGT 1139
 QY 2372 ATGATGTGAGTTGACTGTGAAGAAAGCTGAGCACTGAAGAAATGATCTTTGAAC 2431
 Db 1138 ATGATGTGA-AGTTCGACTGTAAGAAAGCTGAGTCTGAAGAAATGATCTTTGAAC 1080
 QY 2432 TGTGGTGTGGAGAAAGACTCTTTGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAG-TCC 2490
 Db 1079 TGTGGTGTGGAGAAAGACTCTTTGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAGTTC 1020
 QY 2491 ATTCTGAAGAGATCAGCCCTGGGATTTCTTGGAGGAATGATGCTAAAGCTGAAACTC 2550
 Db 1019 ATCCTAAGGAAATTCAGTCTGGGTGTTCAATTTGAGGAGTGTATTTGAAGCTGAAACTC 960
 QY 2551 CAGTACTTTGGCCACCTGTGATCAGAGAGCTGACTCACTGGAAGAGCCCT-GATGCTGGG 2609
 Db 959 CAATACTTTGGCCACCTGTGATGCGAGAGCTGACTCACTTTGAAGAGCCCTGGATGCTGG 900
 QY 2610 AGGATTTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGAGATGGCTGATGCATCAC 2669
 Db 899 AAGATTTGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGATGAGATGGCTGATGCATCAC 840
 QY 2670 TGACTCGATGAGTGTGCTGAGTGAATCTCTGGAGTGGTGTGATGAGAGAGAGAGAGAG 2729
 Db 839 CAACACAAATGACATGGGTTTGGTAAAGTCCGGAGTGGTGTGATGAGAGAGAGAGAG 780
 QY 2730 TCTGCGGCGATTCATGGGTGCAAAAGAGTTGGACACGACTGAGCAACTGAACTGAAC 2789
 Db 779 GTGTGCTGCAAGTTTCATGGGTTGCAAGAGTGCAGATGACTGAGGAGCTGTACTGAAC 720
 QY 2790 GAAC 2793
 Db 719 GAAC 716
 RESULT 27
 ID ABX41993
 XX ABX41993 standard; cDNA; 436 BP.
 XX ABX41993;
 XX 20-FEB-2003 (first entry)
 DT

XX Bovine EST associated with lactation/muscle/fat deposition #7158.
 XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX Bos Taurus.
 XX US2002137139-A1.
 XX 26-SEP-2002.
 XX 24-SEP-2001; 2001US-00960352.
 XX 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX Claim 2; SEQ ID NO 7158; 245pp; English.
 XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 XX Sequence 436 BP; 154 A; 81 C; 107 G; 94 T; 0 U; 0 Other;
 SQ
 Query Match 9.7%; Score 347.4; DB 1; Length 436;
 Best Local Similarity 92.8%; Pred. No. 4.7e+02;
 Matches 385; Conservative 0; Mismatches 26; Indels 4; Gaps 2;
 QY 1931 GAGACATCATGAGAAAGCTGGGCTGGAAGACACAGCTGGAATCAAGATTGCCGGA 1990
 Db 24 GATACATCATGAAACCCCTGGGCTGGAGGAAGTAAAGCTGGAATCAAGATTGCCGGA 83
 QY 1991 GAAATAGCAATAACCTTCAGATATGCAGATATGACCATCCCTTATGGCAGAAAGTGAAGG 2050
 Db 84 GAAATATCAATAGGCTTGATGCAGATGACACACCCCTTATGGCAGAAAGTGAAGG 143

QY 2051 AACTAAAAACCTCTTGATGAGTGAAGAGGAGAGTGAAAAAGTTGGCTTAAAGCTCA 2110
 DB 144 AACTAAAGAGCTCTTGATGAAGGTGAAGAGGAGAGTAAAAAGTTGGCTTAAAGCTCA 203
 QY 2111 ACATTGAGAAACGAGATCATGTCATGTCCTCCATCACTTCTGGAATAGATGGGG 2170
 DB 204 ACATTGAGAAACGAGATCATGTCATGTCCTCCATCACTTCTGGAATAGATGGGG 263
 QY 2171 AAACAGTGAACACAGTGTCAAGCTTTATTTTGGGGGGCTCCAAAATCACTGAGATGG 2230
 DB 264 AAACAGTGAACACAGTGTCAAGCTTTATTTGGGGGGCTCCAAAATCACTGAGATGG 321
 QY 2231 TGACTGCAACCTGAAATTAAGACACTTACTCTTGGAGAAAGTTA--ACCAACCT 2288
 DB 322 TGATTGCAACCTGAAATTAAGACACTTACTCTTGGAGAAAGTTAAGCAACCT 381
 QY 2289 AGATAGCATATTGAAAGAGAGACATTAACCTTGCCAAAGGCCCATCTAGTC 2343
 DB 382 AGATAGCATATTGAAAGAGAGACATTAACCTTGCCAAAGGCCCATCTAGTC 436

RESULT 28

ID ABQ75995 standard; DNA; 31412 BP.
 XX

AC ABQ75995;

DT 05-NOV-2002 (first entry)

XX Sheep PrP gene nucleic acid sequence.

DE Sheep; PrP; prion; transmissible spongiform encephalopathy; TSE; scrapie;
 KW breeding; selection; gene; ds.

XX Ovis aries.

XX GB2371048-A.

XX 17-JUL-2002.

XX 10-JAN-2001; 2001GB-00000702.

XX 10-JAN-2001; 2001GB-00000702.

XX (UYVO-) UNIV YORK.

XX Ross J, Bowles D;

XX WPI; 2002-577556/62.

PT Assay for comparing nucleic acid sequence with a reference, useful for
 PT determining prion gene alleles, comprises determining the profile of
 PT heteroduplexes.

PS Claim 2 (i) (a); Fig 1; 36pp; English.

CC The invention relates to an assay for determining allelic variations in
 CC prion protein genes, involving comparing a nucleic acid sequence with a
 CC reference. The method is used for determining the PrP allele in an animal
 CC (or human) in order to assess susceptibility to transmissible spongiform
 CC encephalopathies (TSE), most especially scrapie in sheep. This assists in
 CC the selection, for breeding, of animals having the alleles known to be
 CC associated with lowest risk. The method is reproducible and rapid,
 CC provides high resolution between allelic variants and is suitable for
 CC processing large numbers of samples. The current sequence represents a
 CC sheep PrP gene, that is used in the context of the invention as a
 CC comparison nucleic acid

XX Sequence 31412 BP; 8889 A; 6517 C; 6464 G; 9542 T; 0 U; 0 Other;

XX Query Match 9.6%; Score 344.199; DB 1; Length 31412;

XX Best Local Similarity 69.7%; Pred. No. 6.9;

Matches 492; Conservative 0; Mismatches 208; Indels 6; Gaps 2;
 QY 2410 TGAAGAATTGATGCTTTTGAACCTGTGGTGTGGAGAGACCTCTTGAGAGTCCCTTGGACT 2469
 DB 17733 TGATTAACAGTAGTGTGTTGAACCTGTGGTGTGGAGAGACCTCTTGAGAGTCCCTTGGACT 17792
 QY 2470 GCAAGAGATCAACACAGTCCATCTCTGAAGAGATCAAGCCCTGGATTTCTTTGGAAGGA 2529
 DB 17793 GCAAGAGATCAACACAGTCCATCTCTGAAGAGATCAAGCCCTGGATTTCTTTGGAAGGA 17852
 QY 2530 ATGATGCTAAAGCTGAAACCTCCAGTACTTTGGCCACCTGATCAAGAGAGTCACTCACTG 2589
 DB 17853 ATGATGCTAAAGCTGAAATTCAGTACTTTGGCCACCTCATGCAAGAGTCACTCACTG 17912
 QY 2590 GAAAGACCTGATCTCTGGAGGATTTGGGGGAGAGAGAGAGGACACAGAGATG 2649
 DB 17913 GAAAGACCTGATCTCTGGAGGATTTGGGGGAGAGAGAGAGGACACAGAGATG 17972
 QY 2650 AGATGGCTGGATGGCATCACTGACTCTCGATGGACCTGAGTCTGGGTGAACCTCTGGAGTTG 2709
 DB 17973 AGATGGCTGGATGGCATCACTGACTCTCGATGGACCTGAGTCTGGGTGAACCTCTGGAGTTG 18032
 QY 2710 GTGATGGACAGGAGGCTCTCTGGGGCGATTCATGGGGTCACAAAGAGTTGACACGA 2769
 DB 18033 TTGATGGACAGGAGGCTCTGGGGCGATTCATGGGGTCACAAAGAGTTGACACGA 18092
 QY 2770 CTGAGCAACTGAACCTGAACCTGACTGAAACCTTAGTAGTTTATATTACTCAGAAA 2829
 DB 18093 CTGAGTGAAGTGAACCTGAACCTGACTGAAACCTTAGTAGTTTATATTACTCAGAAA 18152
 QY 2830 ATAGTAATTCATAGTATCAAAATTAATTCATAGTGTGGTTAAGATAATAGATTTT 2889
 DB 18153 TTGTGTTATTTGGGACATTTCTTCATCTTT-ATTGAATTCCTCAAAAAAATTCACG 18211
 QY 2890 CAAATTTGATTTTATCTTTGATTTTCTCTACTTATTATTTTGGGATTTTAACTATT 2949
 DB 18212 TAGAATTTGATTTGATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 18271
 QY 2950 CTTCAATGACTGTATTTCTTAATTTACTTATTTCTATTTTACT------TTAATTGCACT 3004
 DB 18272 TTATAATACCTGGTGGTTTCCAAACCAAGACCTTTCCATGTTTACAGATCATCTTCTGTGT 18331
 QY 3005 TATTTTATTTGATTTTCTTAATAAATCCAGTCTCTGTTTGTTTTAAAGAGACTTTAAAT 3064
 DB 18332 CTTTAGTAGTGTTCAGTTTCTTTTCCCACTCTACTTTCTTTTGTAGTTAAGTCATCC 18391
 QY 3065 TATTAATTTCTCTTTAGTGTTTTACCAGTCTTTTCCAGGCTACTTCT 3110
 DB 18392 TAGATACCTTACGATGATGATGAGGAGTTGATCTGACTCTTTGT 18437

RESULT 29

ID ABV99702 standard; cDNA; 31412 BP.
 XX

AC ABV99702;

DT 12-FEB-2003 (first entry)

XX Sheep SRE-resistant prion protein cDNA.

DE Transmissible spongiform encephalopathy; neuroprotective; prion protein;
 KW bovine spongiform encephalopathy; transgenic; BSE; bovine; cervid; PrP;
 KW TSE; gene; ss; sheep.

XX Ovis aries.

XX WO200279416-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US0009652.

QY 2401 G----- 2401
 Db 4541 GCTGAGCCCGAAGAAATGATGCTTTTGAACGTGCTGTTGGAGAGACTCTTHSYCNST 4482
 QY 2402 -----CTGAGCACTGAAGAAATTGATGCTTTTGAACGTGCTGTTGGAGAG 2447
 Db 4481 HATGHNNSNASVCTGAGCCCGAAGAAATGATGCTTTTGAACGTGCTGTTGGAGAG 4422
 QY 2448 ACTCTTGAGAGTCCCTTGGACTGCAAGAGATCCCAACGATCCATCTTGAAGGAGATCAG 2507
 Db 4421 ACTCTTGAGAGTCCCTTGGACTGCAAGAGATCCCAACGATCCATCTTGAAGGAGATCAG 4362
 QY 2508 CCTGGGATTTCTTTGGAGGAAATGATGCTAAAGCTGAAATCCAGTACTTTGGCCACCT 2567
 Db 4361 CCTGGGATTTCTTTGGAGGAAATGATGCTAAAGCTGAAATCCAGTACTTAAGCCACCT 4302
 QY 2568 GATCAGAAGCT---GACTCACTGGAAGAAAGCCCTGATGCTGGAGGATTTGGGGCAG 2624
 Db 4301 CATGAGAAGATGACTCATATGGAAGAAAGCTCATGAAAGCTGGAAGCTGGAAGCCAA 4242
 QY 2625 GAGGAGAAGGGGACGACAGAGATGAGATGGCTGGATGGCATCACTGACTCGATGGACGT 2684
 Db 4241 GAGGAGAAGGGGACGACAGAGATGAGATGGCTGGATGGCATCACTGACTCGATGGACAC 4182
 QY 2685 GAGTCTGGTGAATCTCTGAGATTTGATGAGACAGGAGGCTGCTCTGGGGGATTC-C 2743
 Db 4181 GAGTTGAGTGAATCTCTGAGATTTGATGAGACAGGAGGCTGCTCTGGGGGATTC 4122
 QY 2744 ATGGGGTCAAAAGAGTGGGACACGACTGAGCAACTGAACTGAACTGAA 2792
 Db 4121 ATGGGGTCCAAAGAGTGGGACACGACTGAGCAACTGAACTGAACTGAA 4073

RESULT 31

ABX42989
 ID ABX42989 standard; cDNA; 428 BP.
 XX
 AC ABX42989;
 XX
 DT 20-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #8154.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 XX muscle deposition; fat deposition; genome mapping; gene identification;
 XX gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 PN US2002137139-A1.
 XX
 PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-00960352.
 XX
 PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX
 PA (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX
 DR WPI; 2003-110599/10.
 XX
 PT New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX
 PS Claim 2; SEQ ID NO 8154; 245bp; English.

XX
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC : (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Notes: The
 CC present sequence was not shown in the USPTO web site:
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docID=20020137139
 XX
 SQ Sequence 428 BP; 125 A; 74 C; 123 G; 105 T; 0 U; 1 Other;

Query Match 9.4%; Score 335.8; DB 1; Length 428;

Best Local Similarity 86.4%; Pred. No. 5.2e+02;

Matches 370; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 2308 AGAGACATTACCTTCCCAACAAAGCCCATCTAGTCAAGGCTATGTTTCCAGTGGTC 2367
 Db 1 AGAGACATTACCTTCCCAACAAAGGTCGCTAGTAAAGGCTATGTTTCCAGTGGTC 60
 QY 2368 ATGTATGATGTGAGATTGCACTCTGAGAAAGCTGAGCACTCAAGAAATGATGCTTTT 2427
 Db 61 ATGTATGATGTGAGAAATGCACTATATAAGAAATCTGAGCAACCAAGAAATGATGCTTTT 120
 QY 2428 GAACTGTGTTGTGGAGAAAGACTCTTGAGAGTCCCTTGGATCGAAGAGATCAACCCAG 2487
 Db 121 GAACTGTGTTGTGGAGAAAGACTCTTGAGAGTGCCTTGGACTGCAAGAGATCAACCCAG 180
 QY 2488 TCCATTCTGAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTAAGAGCTGAAA 2547
 Db 181 TCCATCTTAAAGAGATCAGTCCCTGAATATTCATTTGGAAGGACTGATGTTGAAGCTGAAG 240
 QY 2548 CTCCAGTACTTTGGCCACCTCTGATCAGAAGAGCTGACTCAGTGGAAAAAGACCCCTGATCTG 2607
 Db 241 CTCCAGTACTTTGGCCACCTCTGATGTAAGAGCTGACTCATTGAAAAAGCCATGATCTG 300
 QY 2608 GGAGGATTTGGGGCAGGAGAGAGGGGACGACAGAGATGAGATGGCTGGATGGCATC 2667
 Db 301 GGAAAGATTAAGTGAATGAGGAGAAAGGGATGACAGAAAGATGATGTTGGATGGCATC 360
 QY 2668 ACTGACTCGATGGACGTGAGTCTGGGTGAATCTCTGGAGTTGGTGTGATGAGCAGGAGGCC 2727
 Db 361 ACCGACTCAATGGACATGGTGGTGGTTCGATCCAGAAAGTTGGTGTGATGAGCAGGAGCC 420
 QY 2728 TGTCTCTGC 2735
 Db 421 TGGCGTGC 428

RESULT 32

ABX36143/c
 ID ABX36143 standard; cDNA; 406 BP.
 XX
 AC ABX36143;
 XX

DT XX 20-FEB-2003 (first entry)
 DE XX Bovine EST associated with lactation/muscle/fat deposition #1308.
 XX XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 XX gene analysis; cattle breeding.
 XX XX
 OS Bos Taurus.
 XX XX
 PN US2002137139-A1.
 XX XX
 PD 26-SEP-2002.
 XX XX
 PF 24-SEP-2001; 2001US-00960352.
 XX XX
 PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX XX
 PA (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TACN/) TAO N.
 PA (WARR/) WARREN W C.
 XX XX
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 DR XX
 PT New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX XX
 PS Claim 2; SEQ ID NO 1308; 245bp; English.
 XX XX
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridization between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docID=20020137139
 XX XX
 SQ Sequence 406 BP; 95 A; 138 C; 77 G; 95 T; 0 U; 1 Other;
 XX XX
 Query Match 9.3%; Score 333; DB 1; Length 406;
 Best Local Similarity 91.9%; Pred. No. 5.6e+02;
 Matches 351; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
 QY 2405 AGCACTGAAGATTGATGCTTTTGAACCTGTGGTGTGGAGAGACTCTTGAGAGTCCCTT 2464
 DB 406 AGTCCCGAAGATTGATGCTTTTGAACCTGTGGTGTGGAGAGACTCTTGAGTGTCCCTT 347
 QY 2465 GGACTCGAAGAGATCCACCACTGCTTCTGAGGAGATCAGCCCTGGGATTTCTTGG 2524

Db 346 GGACTGAAAGAGTTCCAAACCAAGTCCATTCTGAGGAGATCAGCCCTGGGATTTCTTTGG 287
 QY 2525 AAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTTGGCCACCTGATCAGAGAGCTGACT 2584
 Db 286 AGGAATGATGCTGAAAGATGAAACTCCAGTACTTTTGGCCACCTCATGCGAAGAGCTGGCT 227
 QY 2585 CACTGAAAAGACCCCTGATGCTGGGAGGATTTGGGGCAGGAGGAGAGAGGGGACGACAGA 2644
 Db 226 CATTGAAAAGACTCTGATGCTGGGAGGATTTGGGGCAGGAGGAGAGAGGGGACGCCAGA 167
 QY 2645 GGATGAGATGGCTGGATGGCATCACTGACTCGATGACCTGAGTCTGGTGAACCTCTTGG 2704
 Db 166 GGATGAGATGGCTGGATGGCATCGCTGACTCGATGGACGTGAGTCTGAGTGAACCTCCGG 107
 QY 2705 AGTTGCTGATGACAGAGGAGGCTGCTCTGCGGCGGATTCATGCGGTCAAAAAGATTGGA 2764
 Db 106 AGTTGCTGATGACAGAGGAGGCTTGGCTGCTGTGATTCTATGGGTCGCAAGAGTCCGA 47
 QY 2765 CACGACTGAGCAACTGA 2786
 Db 46 TAGGGCTGAGCGACTGA 25
 RESULT 33
 AAH21026
 ID AAH21026 standard; DNA; 1529 BP.
 XX XX
 AC AAH21026;
 XX XX
 DT 31-AUG-2001 (first entry)
 XX XX
 DE Bovine-derived DNA fragment contig1 487_67.
 XX XX
 KW Bovine; digital DNA signature; breeding; animal product origin;
 KW identification; genetic association; breed; population; race; ds.
 XX XX
 OS Bos taurus.
 XX XX
 PN DE19959751-A1.
 XX XX
 PD 13-JUN-2001.
 XX XX
 PF 11-DEC-1999; 99DE-01059751.
 XX XX
 PR 11-DEC-1999; 99DE-01059751.
 XX XX
 PA (FRIE/) FRIES H R.
 PA (DURS/) DURSTEWITZ G.
 XX XX
 PI Fries HR, Durstewitz G;
 XX WPI; 2001-376309/40.
 XX XX
 PT New bovine genomic DNA sequences, useful for establishing genetic
 PT signatures, e.g. for breeding control, contain specific variable
 PT positions.
 XX XX
 PS Claim 1; Page 23-24; 26pp; German.
 XX XX
 CC This invention describes bovine DNA sequences (A) which are used in a
 CC method to establish a digital, standardized DNA signature. DNA signatures
 CC established from (A) are used to monitor breeding; to determine origin of
 CC animal products; to identify individual animals; to study genetic
 CC association and to establish signatures that are specific at the level of
 CC breed, population or race. The method is based on individual base
 CC exchanges in DNA, and these are inherited more stably (by an order of
 CC magnitude) than conventional markers. It is suitable for population-wide
 CC studies (as a high throughput test) and the presence of unequivocal
 CC sequences flanking the variable position provides a built-in
 CC standardization feature
 XX SQ Sequence 1529 BP; 589 A; 299 C; 239 G; 358 T; 0 U; 44 Other;

XX	Query Match	9.1%;	Score 326.3;	DB 1;	Length 1529;
PA	Best Local Similarity	73.7%;	Pred. No. 1.6e+02;		
PA	Matches 452;	Conservative	0;	Mismatches 152;	Indels 9; Gaps 3;
PA	(BYAT/) BYATT J C.				
PA	(MATH/) MATHIALAGAN N.				
PA	(TAON/) TAO N.				
PA	(WARR/) WARREN W C.				
XX	Byatt JC, Mathialagan N, Tao N, Warren WC;				
XX	WPI; 2003-110599/10.				
XX	New nucleic acid associated with lactation, and muscle and fat				
XX	deposition, useful for genome mapping, gene identification and analysis,				
PT	cattle breeding, or for genetically improving cattle.				
PT	Claim 2; SEQ ID NO 8311; 245pp; English.				
XX	The invention relates to a purified nucleic acid molecule associated with				
XX	lactation or muscle and fat deposition (designated LMPD), derived from				
CC	cattle, and the LMPD nucleic acid can specifically hybridize to a second				
CC	nucleic acid molecule comprising any of 15112 nucleotide sequences,				
CC	appearing as ABX4836-ABX4947, or complements of them. Also included are				
CC	; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic				
CC	acid linked to a promoter and a 3' non-translated sequence that				
CC	functions in the cell to cause termination of transcription and addition				
CC	of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and				
CC	(2) determining a level or pattern of a molecule in a bovine cell or				
CC	tissue comprising: (a) incubating a marker nucleic acid (comprising any				
CC	of the 15112 nucleic acid sequences or its complement or fragment) with a				
CC	complementary nucleic acid molecule obtained from the bovine cell or				
CC	tissue, where hybridisation between the marker nucleic acid and the				
CC	complementary nucleic acid permits the detection of the molecule; and (b)				
CC	detecting the level or pattern of the complementary nucleic acid, where				
CC	the detection of the complementary nucleic acid is predictive of the				
CC	level or pattern of the molecule. The LMPD nucleic acid is used for				
CC	determining a level or pattern of a molecule in a bovine cell or tissue.				
CC	It is useful for genome mapping, gene identification and analysis, cattle				
CC	breeding, preparation of constructs for use in cattle gene expression, or				
CC	for genetically improving cattle. The present sequence is one of the				
CC	15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The				
CC	present sequence was not shown in the specification but was obtained in				
CC	electronic format from the USPTO web site:				
CC	seqdata.uspto.gov/sequence.html?DocID=20020137139				
XX	Sequence 423 BP; 93 A; 128 C; 72 G; 130 T; 0 U; 0 Other;				
SQ	Query Match 9.1%; Score 324; DB 1; Length 423;				
	Best Local Similarity 86.7%; Pred. No. 5.6e+02;				
	Matches 357; Conservative 0; Mismatches 55; Indels 0; Gaps 0;				
QY	2380 GAGATTGCACTGTGAAGAAAGCTGAGCAGTGAAGAAATGTGCTTTGAACTGTGGTGT	2439			
Db	423 GAGATTGCACTGTGAAGAAAGCTGAGCAGTGAAGAAATGTGCTTTGAACTGTGGTGT	364			
QY	2440 TGGAGAAGACTCTTGAGAGTCCCTTTGGACTGCAAGAGAGATCCAAAGTCCATTCTGAAG	2499			
Db	363 TGGAGAAGACTCTTGAGAGTCCCTTTGGACTGCAAGAGAGATCCAAAGTCCATTCTGAAG	304			
QY	2500 GAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCTTAAGCTGAAATCTCAGTACTTT	2559			
Db	303 GAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCTTAAGCTGAAATCTCAGTACTTT	244			
QY	2560 GGCCACCTGATCAGAAAGAGCTGACTCACTGGAAGAAAGCCCTGATGCTGGAGGGATTTGG	2619			
Db	243 GGTCACTGATGGGAAAAGAACTGATGCTGAAAAGACCTGATGCTGGGAAAAGATTGAA	184			
QY	2620 GGCAGGAGGAGAAAGGGGACGACAGAGGATGAGATGGCTGGATGGCATCACTGACTCGATG	2679			
Db	183 GGCAGGAGGAGAAAGGGGACGACAGAGGATGAGATGGCTGGATGGCATCACTGACTCTATG	124			
QY	2680 GACGTGAGTCTGGGTGAATCTCTCTGAGTTGGTGGTGGAGAGGAGGCTGTCTCTGGGGG	2739			
Db	123 AACTTGAGTTTGAATAAGCCCGAGGAGTTGGTGGTGGAGAGGAGGAGGAGGAGGAGGAG	64			

QY 2740 ATTCATGGGGTCAAAAGAGTTGGACAGACTGACCACTGAACCTGAACCTGA 2791
 Db 63 GTCCATGGGGTGGCAAGAGTTGGACACAACTAAGTGAAGTGAACCTGAACCTGA 12

RESULT 35

ABX48669
 ID ABX48669 standard; cDNA; 410 BP.

XX AC ABX48669;

XX DT 21-FEB-2003 (first entry)

XX DE Bovine EST associated with lactation/muscle/fat deposition. #13834.

XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;

XX KW muscle deposition; fat deposition; genome mapping; gene identification;

XX KW gene analysis; cattle breeding.

XX OS Bos Taurus.

XX FN US2002137139-A1.

XX PD 26-SEP-2002.

XX PF 24-SEP-2001; 2001US-00960352.

XX PR 12-JAN-1999; 99US-0115707P.

XX PR 11-JAN-2000; 2000US-00480902.

XX PA (BYAT/) BYATT J C.

XX PA (MATH/) MATHIALAGAN N.

XX PA (TAON/) TAO N.

XX PA (WARR/) WARREN W C.

XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX DR WPI; 2003-110599/10.

XX PT New nucleic acid associated with lactation, and muscle and fat

XX PT deposition, useful for genome mapping, gene identification and analysis,

XX PT cattle breeding, or for genetically improving cattle.

XX PS Claim 2; SEQ ID NO 13834; 245pp; English.

XX CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?docID=20020137139

XX SQ Sequence 410 BP; 132 A; 67 C; 104 G; 106 T; 0 U; 1 Other;

Query Match 9.0%; Score 321.6; DB 1; Length 410;
 Best Local Similarity 88.9%; Pred. No. 5.9e+02;
 Matches 368; Conservative 0; Mismatches 40; Indels 6; Gaps 2;

QY 2242 ATGAAATTAAGACACACTTACTCTTGGAGAAAGTTA--ACCACCTAGATAGCATAT 2299
 Db 1 ATGAAATTAAGAGATGCTTACTCTTGGAGAAAGTTATGACCAACCTAGATAGCATAT 60
 QY 2300 TGAAGAAGCAGACACATTACCTTTGCCAACAAAGCCCCCATCTAGTCAAGGCTATGGTTTTTC 2359
 Db 61 TCAAGAAGCAGACACATTACTTTGCCAACAAAGGTCCTGTAGTCAAGGCTATGGTTTTTC 120
 QY 2360 CAGTGGTCATGTATGATGTGAGAGTTGGAGCTGTGAAGAAAGCTGAGCACTGAAGAATTG 2419
 Db 121 CAGTAGTCATGTACGATGTGAGAGTTGGAGCTGTGAAGAAAGCTGAGCACTGAAGAATTG 180
 QY 2420 ATGCTTTTGAAGTGTGGTGGAGAGACTCTTGAGAGTCCCTTGGAGTCCCAAGGAGAT 2479
 Db 181 ATGCTTTTGAAGTGTGGTGGAGAGACTCTTGAGAGTCCCTTGGAGTCCCAAGGAGAT 240
 QY 2480 CCAACCAAGTCCATCTTGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATCTAA 2539
 Db 241 CCAACCAAGTCTATTCTAAAGGAGATCAGTCCCTG---TTCITTTGGAAGGACTGAAGCTAA 296
 QY 2540 AGCTGAAAGTCCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCCTGGAAGAGACCC 2599
 Db 297 AGCTGAAAGTCAAATTTCTTTGGCCACCTGATCAGAGAGAGTTGACTCATTTGGAAGAGACTT 356
 QY 2600 TGATGCTGGAGGGATTGGGGGACGAGGAGAGAGAGGAGGAGGAGATGATGATGAT 2653
 Db 357 TGATGCTTGGAAAGACTGAAGGCTGGAAGAGAGAGGAGATGACAGANGATGATGAT 410

RESULT 36

ABX40334/c
 ID ABX40334 standard; cDNA; 423 BP.

XX AC ABX40334;

XX DT 20-FEB-2003 (first entry)

XX DE Bovine EST associated with lactation/muscle/fat deposition #5499.

XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;

XX KW muscle deposition; fat deposition; genome mapping; gene identification;

XX KW gene analysis; cattle breeding.

XX OS Bos Taurus.

XX PN US2002137139-A1.

XX PD 26-SEP-2002.

XX PF 24-SEP-2001; 2001US-00960352.

XX PR 12-JAN-1999; 99US-0115707P.

XX PR 11-JAN-2000; 2000US-00480902.

XX PA (BYAT/) BYATT J C.

XX PA (MATH/) MATHIALAGAN N.

XX PA (TAON/) TAO N.

XX PA (WARR/) WARREN W C.

XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;

XX DR WPI; 2003-110599/10.

XX PT New nucleic acid associated with lactation, and muscle and fat

XX PT deposition, useful for genome mapping, gene identification and analysis,

XX PT cattle breeding, or for genetically improving cattle.

XX PS Claim 2; SEQ ID NO 5499; 245pp; English.

XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridization between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 423 BP; 93 A; 128 C; 72 G; 130 T; 0 U; 0 Other;
Query Match 9.0%; Score 320.8; DB 1; Length 423;
Best Local Similarity 86.2%; Pred. No. 5.7e+02;
Matches 355; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
Qy 2380 GAGAGTTGGACTGTGAAGAAGCTGACACTGAGAAATGATGCTTTTGAACCTGTGTGT 2439
Db 423 GAGAGTTGGACTGTGAAGAAGCTGAGCGCCAAAGAAATGATGCTTCTGAACCTGTGTGT 364
Qy 2440 TGGAGAGACTCTGAGAGTCCCTTGGAGGAGATGCTGACGAGATCCACAGTCCATCTCGAG 2499
Db 363 TGGAGAGACTCTTGAAGTCCCTTGGAGGAGATCCACAGTCCATCTCGAG 304
Qy 2500 GAGATCAGCCCTGGGATTTCTTGGAAAGGAATGATGCTAAGCTGAAACCTCCAGTACTTT 2559
Db 303 GAGATCAGCCCTGGGATTTCTTGGAAAGGAATGATGCTAAGCTGAAACCTCCAGTACTTT 244
Qy 2560 GGCACCTGTATCAGAGAGCTGACTACTGAAAGACCTGATGCTGGAGGAGATTGG 2619
Db 243 GGTCACTGTATGGGAAAACCTGACTGATCTGAAAAGACCTGATGCTGGGAAAGATTGAA 184
Qy 2620 GGCAGGAGGAAAGGAGACAGAGGATGATGGCTGGATGGCATCACTGACTCGATG 2679
Db 183 GGCAGGAGGAAAGGAGACAGAGGATGATGGCTGGATGGCATCACTGACTCGATG 124
Qy 2680 GACGTGAGTCTGGGTGAACCTCCTGGAGTTGTGTGATGACAGGAGGCTGTCTCTGGGGG 2739
Db 123 AAC*TGAGTTTGAATTAAGCCCGAGTGGTGTGATGACAGGAGGAGCATGACTACTGCA 64
Qy 2740 ATTATGGGTCACAAAGATTGACAGCTGACGACTGAACTGA 2791
Db 63 GTCCATGGGTTGCAAGAGTTGGACACAACTAAGTACTGACTGA 12
RESULT 37
ID ABX49143/C
XX ABX49143 standard; cDNA; 409 BP.
XX AC ABX49143;
XX XX
XX DT 21-FEB-2003 (first entry)
XX DE Bovine EST associated with lactation/muscle/fat deposition #14308.
XX

KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
OS Bos Taurus.
XX US2002137139-A1.
PN 26-SEP-2002.
XX 24-SEP-2001; 2001US-00960352.
XX 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
PI WPI; 2003-110599/10.
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX Claim 2; SEQ ID NO 14308; 245pp; English.
PS The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridization between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 409 BP; 90 A; 97 C; 84 G; 137 T; 0 U; 1 Other;
Query Match 9.0%; Score 320.4; DB 1; Length 409;
Best Local Similarity 89.3%; Pred. No. 6e+02; Indels 2; Gaps 2;
Matches 366; Conservative 0; Mismatches 42; Indels 2; Gaps 2;
Qy 1748 AAACGTGGAAAATTTCTGAAGGATGGGAATACAGACCCACCTGACCTCTTGAAA 1807
Db 409 AAACGTGGACAATTTCTGAAGAGATGGGAATACAGACCCACCTGATCTGCTTGA 350
Qy 1808 AATTTGTATGCGAGTCAGGACAGACAGTATAGAACTGGACATGGACACAGCTGTTTC 1867
Db 349 AATCTGTATGCGAGTCAGGACAGACAGTATAGAACTGGACATGGACACAGCTGTTTC 290
Qy 1868 CAAGTAGGAAAAGGAGTATGTCAGGCTGTATATGTCACCCGGCTTGTAACTTCTAT 1927

Db 289 CAATAGGACAGGAATCCGTCAAGTAGTAGTATATGTACCCCTGCTATTATTAACCTTATAT 230
QY 1928 GCAGAG-ACATCATGAGAAACGCTGGGTGGAAGAAGCACAAGCTGGAATCAAGATTGCC 1986
Db 229 GCAGAGTCCATGATGAGAAACCTCTGGGCTGGAAGAAACACAAGCTGGAATCAAGATTGCC 170
QY 1987 GGGAGAAATAGCAATAACCTCAGATATGCAGATGATACCAACCTTATGGCAGAAAGTGAA 2046
Db 169 GGGAGAAATATCAATAACCTCAGATATGCAGATGATACCAACCTTATGGCAGAAAGTGAA 110
QY 2047 GAGGAACATAAAGCCTCTTCATGAAAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAG 2106
Db 109 GAGGAACATAAAGCCTCTTCATGAAAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 50
QY 2107 CTCACATTCAGAAACAGCAATGATGCAATGCTGCTGCTCCATCATTCTATG 2156
Db 49 CTCACCTTCAGAAATG-NGATCATGCGATCCGTCCTCCGTTCCATG 1

RESULT 38
AAK56452/c
ID AAK56452 standard; cDNA; 529 BP.
XX
AC AAK56452;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1512.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytosolic; gene therapy; vaccine; metastasis; ss.
XX Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180625P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246521P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.

PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 03-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254057P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-483426/52.
DR P-PSDB; AAM83671.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Claim 1; SEQ ID NO 1512; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 529 BP; 121 A; 124 C; 107 G; 171 T; 0 U; 6 Other;
SQ
Query Match 8.9%; Score 316.4; DB 1; Length 529;
Best Local Similarity 88.8%; Pred. NO. 4.7e+02;
Matches 372; Conservative 1; Mismatches 38; Indels 8; Gaps 3;
QY 604 GTCAACAGAGATGGCAGAGTGAATGTCACATCTTCTAGGAACAGCGAAC-TAAATG 662
DB 413 GTCAACAGAGATGGCAGAGTGAATGTCACATCTTCTAGGAACAGCGAAC-TTAAATG 354
QY 663 GACTGGAATGGTGAATTTAACTCAGATGACCATTTATCTACTACTCGCGGCGAGGATC 722
|||||

Db 353 GACTGGAATGGTGAATTTAACTCAGATAACCAATTATATCTACTACCGTGGCGAGGAATC 294
QY 723 CCTCAGAGAAATGGAGTAGCCATCATGCTCAACAAAGAGTCCGAAATGCAGTACTTGG 782
Db 293 CCTTAGAAGAAACAGAGTAGCCATCATGCTCAACAAAGAGTCTGAAATGCAGTACTTGG 234
QY 783 ATGAGTCTCAAAAACGACAGAAATGATCTCTGTTTGTTCACAGGCAAAACCATTCATAT 842
Db 233 ATGCAATCTCAAAAACGAGAAATGA-----CTGGTTTCCAGGCAAAACCATTCATAT 180
QY 843 CACAGTAATCAAGTCTATGCTCCCAACCAAGTAACTGCTGAAGAGCTGAAGTTGAACGTC 902
Db 179 CACAGTAATCCAGCCCTATGCTCCCAACCAAGTAACTGCTGAAGAGCTGAAGTTGAATGTT 120
QY 903 CTATGAAGACCTACAAGACCTTTTAGAAGTAACTAACACCCCAAAA-AGATGTCTCTTCATTA 961
Db 119 CTATGAAGACCTACAAGACCTTTTAGAAGTAACTAACACCCCAAAAAGATGTCTTCGTTA 60
QY 962 TAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGGAGTAAACAGGCAAAATTT 1020
Db 59 TGGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACATCTGGGGTAAACAGAGAAATTT 1
RESULT 39
ABX44380
ID ABX44380 standard; cDNA; 425 BP.
XX
AC ABX44380;
XX
DT 21-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #9545.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-00960352.
XX
PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARRE/) WARREN W C.
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 9545; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMPD), derived from
CC cattle, and the LMPD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 1512 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
CC acid linked to a promoter and a 3' non-translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or

CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 SQ Sequence 425 BP; 144 A; 70 C; 100 G; 111 T; 0 U; 0 Other;

Query Match 8.8%; Score 313.4; DB 1; Length 425;
 Best Local Similarity 87.4%; Pred. No. 6e+02;
 Matches 375; Conservative 0; Mismatches 46; Indels 8; Gaps 3;
 QY 2073 GGTGAAGAGAGAGTGAAGAAGTGGCTTAAGCTCAACATTCAGAAACGAGATCAT 2132
 Db 3 GGTGAAGAGAGAGTGAAGAAGTGGCTTAATCTCAACATTCAGAAACGAGATCAT 62
 QY 2133 GGCATCTGGTCCCATCACTTCATCGGAATAAGATGGGAAACAGTGGAAACAGTGTGAGA 2192
 Db 63 GGCAGCTGATCCCATTAATTCATGGCAATAGATGGGGATACAAATGGAACAGTGGAGAGA 122
 QY 2193 CTTTATTTTGGGGGCTCCAAATCACTGCAGATGGTGAAGTGGTCAAGCCATGAATTA 2252
 Db 123 CTTTATGTTTT--TGGGCTCCAAATCACTGCAGATGGTGAAGTGGTCAAGCCATGAATTA 180
 QY 2253 AGACATCTACTCTTGGGAAGAAAGTTA--ACCAACCTAGATAGCATATTTGAAGAGCAGA 2310
 Db 181 TGACTCTTCTTGGGAAGAAAGTATGACCGAGCTAGACAGCATATTTAAAGAGCAGA 240
 QY 2311 GACATTAACCTTGCACAAAGCCCATCTAGTCAAGCTATGGTTTTCAGTGGTCAATG 2370
 Db 241 AACATTAATTTACCAACAAAGTCCATCTATTCAAAGCTA---TTTCTAGTAGTCAATG 296
 QY 2371 TATGGATGTGAGAGTTGGACTGTGAAGAAAGCTGAGCACTGAAGAAATTGATGCTTTTGA 2430
 Db 297 TATGTTGTGAGAGTTGGACTATTAAGAAAGCTGAGTCTGAAGAACTGATGGTTTGA 356
 QY 2431 CTGTGGTGTGGAGAGACTCTTGAGAGTCCCTTGGACTCCAAAGGAGATCCAAACCAAGTCC 2490
 Db 357 CTGGGTGTGGAGAGACTCTGAGAGTCCCTTGAAGTCCCTTGAAGTCCAAAGGAGATCCAAACCAAGTCA 416
 QY 2491 ATTCCTGAAG 2499
 Db 417 ATCCTAAG 425

RESULT 40
 AAH21024/C
 ID AAH21024 standard; DNA; 1435 BP.
 XX
 AC AAH21024;
 XX
 DT 31-AUG-2001 (first entry)
 XX
 DE Bovine-derived DNA fragment contig1 422_19.
 XX
 KW Bovine; digital DNA signature; breeding; animal product origin;
 XX identification; genetic association; breed; population; race; ds.
 XX
 OS Bos taurus.
 XX
 PN DE19959751-A1.
 XX

PD 13-JUN-2001.
 XX
 PF 11-DEC-1999; 99DE-01059751.
 XX
 PR 11-DEC-1999; 99DE-01059751.
 XX
 PA (FRIE/) FRIES H R.
 PA (DURS/) DURSTEWITZ G.
 XX
 PI Fries HR, Durstewitz G;
 XX WPI; 2001-376309/40.
 DR
 XX New bovine genomic DNA sequences, useful for establishing genetic
 PT signatures, e.g. for breeding control, contain specific variable
 PT positions.
 XX
 XS Claim 1; Page 22; 26pp; German.
 XX

CC This invention describes bovine DNA sequences (A) which are used in a
 CC method to establish a digital, standardized DNA signature. DNA signatures
 CC established from (A) are used to monitor breeding; to determine origin of
 CC animal products; to identify individual animals; to study genetic
 CC association and to establish signatures that are specific at the level of
 CC breed, population or race. The method is based on individual base
 CC exchanges in DNA, and these are inherited more stably (by an order of
 CC magnitude) than conventional markers. It is suitable for population-wide
 CC studies (as a high throughput test) and the presence of unequivocal
 CC sequences flanking the variable position provides a built-in
 CC standardization feature
 XX

SQ Sequence 1435 BP; 309 A; 301 C; 241 G; 584 T; 0 U; 0 Other;
 Query Match 8.7%; Score 312.4; DB 1; Length 1435;
 Best Local Similarity 74.5%; Pred. No. 1.8e+02;
 Matches 427; Conservative 0; Mismatches 116; Indels 30; Gaps 3;
 QY 1059 CHAGAAATGACCTGCTCATAGCAAAACCCCTCTTCCAAACACACAGAGAACTCTAC 1128
 Db 1079 CAAGAGACATCTATGATCAGAGCAAAACCCCTCTTCCAAATACACAGAGATGACTCTAC 1020
 QY 1129 ACATGGACATCACAGATGGTCAACACCGAATCAGATTGATTATATTCTTTGAGCCAA 1188
 Db 1019 ACATGGACATCACACGTTGGTCAATACCGCAATCAGACTGATTATATTCTTTAGGGCTGA 960
 QY 1189 AGATGGAGAGCTCTATACAGTCAAGCAAAAACACAGACGAGGCTTACTGTGGCTCAGAT 1248
 Db 959 AGATGGACAACTCTATACAGTCAAGCAAAAACACAGACCTAGAG-----TTGACTCAGAT 906
 QY 1249 CATGAACCTCTTATGGCCAAATTCAGACTTAATAATTGAGAAAGTA--GGGAAACCACTAG 1307
 Db 905 TATTAGCACCTCACTGCAAAAAGTCAAGGCTTAATTTGAAGAAAGTAGGGGAAACTACTAG 846
 QY 1308 ATCACTCAGGTAAGACCTTAAATCCAATCCCTTTATGATTATATACAGTGGAAAGTAATAG 1367
 Db 845 GCCATTGAGGATGACCTTAATCAATCTCTTATGATTACACAGTGGGGTGTGAACAG 786
 QY 1368 ATTTAAGGCTTAGATCTGATAGACAGATCACTTAATGAACCTATGACAGAGGTTCTATG- 1426
 Db 785 ATTCAGGGATTAGATCTGTTAGACAGTATGCTCTGAAGAACTATGGAAGTGGAGTTCATGT 726
 QY 1427 -----ACATTGTACAGGACAGGAGTCCAGACCATCCCCCATG 1464
 Db 725 TTGCTCTGACATAACATTTGTACATTTGTACAGAGGCGAGTCAATAAACCATTCCAAAG 666
 QY 1465 GAAAGAAATGCAAAAAGCAAAATGGCTGTCTGGGAGGCTTACAAATAGCTGTGAAA 1524
 Db 665 AAAAAGAAATGCAAGAAGGCAAAAGTGTATCTACAGAGGCTTTACAAATAGCTGAGAA 606
 QY 1525 AGAAGAGAAAGTGAAGAAAGGAAAGAAAGAAAGATATAAAGCATCTGAATGAGAGTTC 1584
 Db 605 AGAAGAGAAAGTGAAGAAAGGAAAGGAAAGATATAAAGGAAAGATATACCTGACTGATATCTGATAGA 546

KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 XX Gene analysis; cattle breeding.
 XX Bos Taurus.
 XX US2002137139-A1.
 XX 26-SEP-2002.
 XX 24-SEP-2001; 2001US-00960352.
 XX 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX Claim 2; SEQ ID NO 11945; 245pp; English.
 XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
 CC nucleic acid molecule comprising any of 1512 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non- translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 1512 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 1512 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX Sequence 447 BP; 162 A; 88 C; 94 G; 103 T; 0 U; 0 Other;

Query Match 8.4%; Score 299.4; DB 1; Length 447;
 Best Local Similarity 84.1%; Pred. No. 6.2e+02;
 Matches 360; Conservative 0; Mismatches 66; Indels 2; Gaps 2;
 QY 1048 CAAGAGCTAATAGATTGTCACAGAAATGCACCTGTCATAGCAACACCTCTTCCAA 1107
 DB 22 CGAAGGCCAGTAAAGTCTTCAACAGAGAAATATATTGGTGATAGC-AGAACCCTTTTCAA 80
 QY 1108 CAACACAAGAGAGCTCTACACATGGACATCCAGATGGTCAACCGGAATCAGATT 1167
 DB 81 CGGCGCAAGATRAAGATTCTACATGTGGACATCATCAGTTAGTCAACACCAAAATCAGATT 140
 QY 1168 GATTATATCTTTGGAGCCAAAGATGGAGAGCTCTATACAGTCAGCAAAAACAGACCA 1227

DB 141 GATTATATCTTTGGAGCAAAAATGG-GGAGCTCTATACAGTCAGCAAAAAGCAAGCTG 199
 QY 1228 GGAGCTTACTGTGGCTCAGATCATGAACCTCTTATTCGCAAAATTCAGACTTAATTTGAAG 1287
 DB 200 GGAGCTGACTGTGCTCTCTGCTCATGAACACCTATTTGCAAAATACAGACTTAAATTTGAAG 259
 QY 1288 AAAGTAGGGAACCACTAGATCACTCAGGTGAAGACCTAAATCCATCCCTTTATGATTAT 1347
 DB 260 AAAGTAGGGAACCACTAGATCACTTATGATATGACCTATATCAACCCCTTACTACTAT 319
 QY 1348 ACAGTGGAAAGTGAGAAATAGATTAAAGGGCTAGATCTGTAGACAGAGTACCTAATGAA 1407
 DB 320 ACAGTGGAAAGTGGGAAATATATTTAAAGGGACTAGATCTGTAGACAGATGCTGTATGAA 379
 QY 1408 CTATGGACAGAGGTTTCATGACATTTGTACAGGACAGGAGATCGAGACCATCCCATCGAA 1467
 DB 380 CTATAGACAGAGGTTTCATGACATTTGTACAGGACAGGAGCAAGACCATCCCAAGATA 439
 QY 1468 AAGAAATG 1475
 DB 440 GAGAAAG 447

RESULT 45
 AAD26490/c
 ID AAD26490 standard; DNA; 815 BP.
 XX
 AC AAD26490;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 XX Sheep alpha (1, 3) galactosyltransferase post-exon 4 DNA #1.
 DE
 XX Sheep; xenotransplantation; Gal-alpha (1,3)Gal; GAL determinant; surgery;
 KW alpha(1,3) galactosyltransferase; alpha (1, 3) GT; drug screening;
 KW gene therapy; ds.
 XX
 OS Ovis sp.
 XX
 XX WO2001188096-A2.
 PN
 XX 22-NOV-2001.
 PD
 XX 14-MAY-2001; 2001WO-US015765.
 PF
 XX 15-MAY-2000; 2000US-0204148P.
 PR
 XX 13-JUN-2000; 2000US-00593316.
 PR
 XX (GERO-) GERON CORP.
 PA
 XX Denning C, Clark J;
 XX
 XX WPI; 2002-089848/12.
 DR
 XX New ovine tissue devoid of antibody-detectable Gal-alpha(1,3)Gal
 PT determinants, for xenotransplantation, and in the treatment of the human
 PT body by surgery or therapy.
 PT
 XX Claim 18; Page 76; 86pp; English.
 PS
 XX The patent discloses immunologically compatible animal tissue, suitable
 CC for xenotransplantation into human patients. The invention also relates
 CC to an ovine tissue devoid of antibody-detectable Gal-alpha (1,3)Gal (GAL)
 CC determinants which are made by alpha(1,3) galactosyltransferase (GT). The
 CC ovine tissue is useful for treatment of human body by surgery or therapy
 CC and in xenotransplantation, by transplanting the ovine tissue into a
 CC mammal having circulating antibody against Gal alpha(1,3)GT determinants.
 CC Polynucleotide constructs of the invention are useful for inactivating an
 CC alpha(1,3) GT gene in an ovine cell. Alpha (1,3) GT sequences are useful
 CC for drug screening and for the production of GAL containing synthetic
 CC oligosaccharides. Sequences of the invention are also useful in gene
 CC therapy. The present sequence is alpha (1,3) GT post-exon 4 DNA

SQ Sequence 815 BP; 217 A; 189 C; 186 G; 221 T; 0 U; 2 Other;

Query Match 8.4%; Score 298.6; DB 1; Length 815;
Best Local Similarity 87.8%; Pred. No. 3.4e+02;
Matches 347; Conservative 0; Mismatches 44; Indels 4; Gaps 2;

QY 2402 CTGAGCACTGAAGAAATTGATGCTTTTGAACCTGTGTGTGTGGAGAAGACTCTTGAGAGTCC 2451
DB 815 CTGAGCCCCGAAGAATCGATGCTTTTGAACCTGTGTGTGTGGAGAAGACTCTTGAGAGTCC 756
QY 2462 CTTGGACTGCAAGGAGATCCACCACTCCATTCTGAAGGAGATCAGCCCTGGGATTTCCTT 2521
DB 755 CTTGGACTGCAAGGAGATCCACCACTCCATTCTGAAGGAGATCAGCCCTGGGATTTCCTT 696
QY 2522 TGGAGGAATGATGCTAAAGCTGAAGTCCAGTACTTTGGCCACCTGATCAGAAGAGCT- 2580
DB 695 TGGAGGAATGATGCTAAAGCTGAAGTCCAGTACTTAAGGCCACCTCATGAGAAGAAGTT 636
QY 2581 --GACTCACTGGAAGACCTGATGCTGGGAGGATTTGGGGCAGGAGGAGGGGAC 2638
DB 635 GACTCATATGGAAGAAAGACTCATGAAGCTGGGAGGATTTGGGGCAGGAGGAGGGGAC 576
QY 2639 GACAGAGGATGAGATGGCTGGATGGCATCACTGACTCCGATGGAGCGTGTGAGTCTGGGTGAAC 2698
DB 575 GACAGAGGATGAGATGGCTGGATGGCAACACCACTCAATGGACACGAGTTTGAGTGAAC 516
QY 2699 TCCTGGAGTTGGTGTGATGGACAGGGAGGCGTGTCTGCGGCGATT- CATGGGCTCACAAG 2757
DB 515 TCTAGGAGTTGGTGTGATGAGACAGGGAGGCGTGTGCGGATTACATGGGGTCGCAAG 456
QY 2758 AGTTGACACGACTGAGCAACTGAAGTGAAGTGA 2792
DB 455 AGTMGGACACGACTGAACGACTGAAGTGAAGTGA 421

Search completed: August 24, 2004, 13:11:25
Job time : 671 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 13:21:40 ; Search time 604 Seconds
(without alignments)
3.966 Million cell updates/sec

Title: US-10-664-775-2
Perfect score: 3572
Sequence: 1 gtcaggaaggcgcagtg.....gcaacaacagcagaagctt 3572

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 45 seqs, 335272 residues

Total number of hits satisfying chosen parameters: 90

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rni2.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	672	18.8	3370	1	US-08-060-822A-5
C 2	672	18.8	3370	1	US-08-060-822A-5
C 3	194.3	5.4	6763	1	US-08-756-506-23
C 4	194.3	5.4	10807	1	US-08-756-506-23
C 5	194.3	5.4	10807	1	US-08-756-506-5
C 6	183.2	5.1	2044	1	US-08-071-601-3
C 7	183.2	5.1	2044	1	US-08-621-100-3
C 8	183.2	5.1	4532	1	US-09-930-377B-1
C 9	164.3	4.6	3317	1	US-09-193-562D-1
C 10	115.6	3.2	19124	1	US-08-487-826B-13
C 11	101	2.8	6669	1	US-10-204-708-6
C 12	100.9	2.8	5862	1	US-10-204-708-6
C 13	100.2	2.8	6070	1	US-10-204-708-10
C 14	100.1	2.8	8961	1	US-10-204-708-80
C 15	97.1003	2.7	5666	1	US-10-204-708-29
C 16	96.1003	2.7	6866	1	US-10-204-708-20
C 17	95.9003	2.7	8607	1	US-10-204-708-72
C 18	95.4003	2.7	6306	1	US-10-204-708-50
C 19	94.7002	2.7	658	1	US-08-998-416-595
C 20	93.9003	2.6	11049	1	US-10-204-708-22
C 21	93.8003	2.6	6326	1	US-10-204-708-57
C 22	92.7003	2.6	10467	1	US-10-204-708-2
C 23	92.4003	2.6	6040	1	US-10-204-708-69
C 24	92.3004	2.6	9347	1	US-10-204-708-35
C 25	91.9999	2.6	7218	1	US-08-232-463-14
C 26	91.5004	2.6	10619	1	US-10-204-708-3
C 27	90.4004	2.5	11049	1	US-10-204-708-23
C 28	89.7004	2.5	8601	1	US-10-204-708-62
C 29	86.9003	2.4	11049	1	US-10-204-708-21
C 30	86.8002	2.4	5455	1	US-10-204-708-33
C 31	85.4004	2.4	636	1	US-08-998-416-1137
C 32	84.5003	2.4	6113	1	US-10-204-708-14
C 33	83.8002	2.3	8093	1	US-10-204-708-31

ALIGNMENTS

RESULT 1

US-08-060-822A-5/c
; Sequence 5, Application US/08060822A
; Patent No. 5432270
; GENERAL INFORMATION:
; APPLICANT: Bevis, Charles L.
; APPLICANT: Diamond, Gill
; APPLICANT: Zaslavoff, Michael
; TITLE OF INVENTION: No. 5432270el Antimicrobial Peptides
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Woodcock Washburn Kurtz Mackiewicz & No. 5432270ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/060,822A
; FILING DATE: 19930511
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/991,200
; FILING DATE: December 15, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca Ralph Gaumond
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURES:
; NAME/KEY: CAAT signal
; LOCATION: 1382..1389
; FEATURE:
; NAME/KEY: TATA signal
; LOCATION: 1444..1449
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1503..1560
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3074..3211
; FEATURE:

Sequence 2, Appl
Sequence 86, Appl
Sequence 55, Appl
Sequence 45, Appl
Sequence 3, Appl
Sequence 38, Appl
Sequence 41, Appl
Sequence 186, Appl
Sequence 71, Appl
Sequence 26, Appl
Sequence 288, Appl
Sequence 39, Appl

34 82.6003 2.3 5852 1 US-07-867-106-2
35 82.4003 2.3 11050 1 US-10-204-708-86
36 81.3004 2.3 11015 1 US-10-204-708-55
37 81.3003 2.3 19233 1 US-10-204-708-45
38 81.0002 2.3 6124 1 US-08-213-419B-3
39 80.4004 2.3 5501 1 US-10-204-708-38
40 80.1003 2.2 8537 1 US-10-204-708-41
41 79.7003 2.2 615 1 US-08-998-416-186
42 78.8004 2.2 8607 1 US-10-204-708-71
43 78.5002 2.2 6583 1 US-10-204-708-26
44 78.3004 2.2 837 1 US-08-998-416-288
45 77.6002 2.2 19513 1 US-10-204-708-39

```

; NAME/KEY: exon
; LOCATION: 1502..1560
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3074..3211
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1..1472
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1561..3073
; FEATURE:
; NAME/KEY: intron
; LOCATION: 3211..3370
; US-08-060-822A-5

Query Match      18.8%; Score 672; DB 1; Length 3370;
Best Local Similarity 89.1%; Pred. No. 8.4e-07;
Matches 778; Conservative 0; Mismatches 85; Indels 10; Gaps 5;

QY 1950 TGGGCTGGAAGACACAGCTGGAATCAAGATTCGCGGAGAAATAGCAATAACCTCAG 2009
Db 939 TTGGCTGGAAGACACAGCTGGAATCAAGATTCGCGGAGAAATGCAATAACCTCAA 880

QY 2010 ATATGAGATGATACACCTTATGCGAGAAAGTGAAGAGAACTAAAAGCCTTTGAT 2069
Db 879 ATATGAGATGATACACCTTATGCGAGAAAGTGAAGAGAACTAAAAGCCTTTGAT 820

QY 2070 GAAGTGAAGAGGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAAGCAAGAT 2129
Db 819 GAAAGTGAAGATGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAAGTGAAT 760

QY 2130 CATGSCATCTGCTCCCATCTTATGCGAAATAGATGGGAAACAGTGAAGACAGTGC 2189
Db 759 AATGCGTATATCCCATCTTATGCGAAATAGATGGGAAACAGTGAAGACAGTGC 700

QY 2190 AGACTTTATTTTGGGGGGCTCAAAATCACTGAGATGTGACTGACGCGATGAAT 2249
Db 699 AGACTTTATTTTGGGGGGCTCAAAATCACTGAGATGTGACTGACGCGATGAAT 643

QY 2250 AAAAGACATTTACTCTTGGGAAGAAAGTTA--ACCACTAGATACATATTGAAAAGC 2307
Db 642 AAAACACAGATTACTCTTGGGAAGAAAGTTATGACCACTAGATACAT--TTGAAAAGC 584

QY 2308 AGAGACATTTACTCTTGGGAAGAAAGTTA--ACCACTAGATACATATTGAAAAGC 2364
Db 583 AGAGATTTACTTTTACCACAAAGGTTCGCTAGTCAAGCCCTATGTTTTTTCAGTG 524

QY 2365 GTCATGATGATGAGTGGAGTTGAGCTGTGAAGAAAGCTGAGCACTGAAGATTTGATGCT 2424
Db 523 GTCATGATGATGAGTGGAGTTGAGCTGTGAAG--AAGCTGAGCGCCAAAGAAATTGATGCT 465

QY 2425 TTTGAACCTGTGTTGCGAGAGACTCTTCAGAGTCCCTTGGACTGCGAGGAGATCCAAC 2484
Db 464 TTTGAACCTGTGTTGCGAGAGACTCTTCAGAGTCCCTTGGACTGCGAGGAGATCCAAC 405

QY 2485 CAGTCCATCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATGCTTAAAGCTG 2544
Db 404 CAGTCCATCTGAAGAGATCAGTCCCTGGGATTTCTTTGGAAGGAATGATGCTTAAAGCTG 345

QY 2545 AAATCCAGTACTTTGGCCACCTGATCAGAAAGTGTACTCTGGAAGAAAGCCTTGATG 2604
Db 344 AAGCTCAATTTGCGGAGGAGTGTGAGAGAGTGTGAGTCAATTTGGAAGAAAGCCTTGATG 285

QY 2605 CTGGAGGAGATTGGGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2664
Db 284 CTGGAGGAGATTGGGGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 225

QY 2665 ATCACTGATGATGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2724
Db 224 ATCAGGAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 165

QY 2725 GCCTGTCTTGGGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2784

```

```

Db 164 TCCTGGCATGCTGCAATTCATGGGTGCGAAAGAGTTGGTACCCTGAGCGACTGAACT 105
QY 2785 GAAGTGAAGTGTACTGAAACCTTAGTAGTTTAT 2817
Db 104 GAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 72

RESULT 2
PCT-US94-05257-5/c
; Sequence 5, Application PC/TUS9405257
; GENERAL INFORMATION:
; APPLICANT: Bevins, Charles L.
; APPLICANT: Diamond, Gill
; APPLICANT: Zaslloff, Michael
; TITLE OF INVENTION: Novel Antimicrobial Peptides
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05257
; FILING DATE: 11-MAY-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/060,822
; FILING DATE: May 11, 1993
; APPLICATION NUMBER: 07/991,200
; FILING DATE: December 15, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca Ralph Gaumond
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0408
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3370 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CAAT signal
; LOCATION: 1382..1389
; FEATURE:
; NAME/KEY: TATA signal
; LOCATION: 1444..1449
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1503..1560
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3074..3211
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1502..1560
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3074..3211
; NAME/KEY: intron
; LOCATION: 1..1472

```

FEATURE:
NAME/KEY: intron
LOCATION: 1561...3073
FEATURE:
NAME/KEY: intron
LOCATION: 3211...3370
PCT-US94-05257-5

Query Match
Best Local Similarity 89.1%; Score 672; DB 1; Length 3370;
Matches 778; Conservative 0; Mismatches 85; Indels 10; Gaps 5;

Qy 1950 TGGCTGGAGACACAGCTGCAATCAAGATTGCGGAGAAATAGCAATTAACCTCAG 2009
Db 939 TTGCTGGAGACACAGCTGCAATCAAGATTGCGGAGAAATAGCAATTAACCTCAA 880

Qy 2010 ATATGTCAGATCATACCACTTATGCGAGAAAGTGAAGAGAACTAAAAAGCCTCTTGAT 2069
Db 879 ATATGTCAGATCATACCACTTATGCGAGAAAGTGAAGAGAACTAAAAAGCCTCTTGAT 820

Qy 2070 GAAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAGAT 2129
Db 819 GAAAGTGAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACGAAGAT 760

Qy 2130 CATGCACTCTGCTCCCTCACTTCTATGGAATAGATGGGAAACAGTGAACAGTGC 2189
Db 759 AATGGCTATGATCCATCACTTCTATGGAATAGATGGGAAACAGTGAACAGTGC 700

Qy 2190 AGACTTTATTTTTGGGGGCTCCAAATCACTCCAGATGTGACTGCAGCCATGAATTT 2249
Db 699 AGACTTTATTTTTGGGGGCTCCAAATCACTCCAGATGTGACTGCAGCCATGAATTT 643

Qy 2250 AAAAGACACTTACTCTTGGAGAAAGTTA--ACCACTTAGATAGCATATTTGAAAGC 2307
Db 642 AAAAGACACTTACTCTTGGAGAAAGTTATGACCACTTAGATAGCAT--TTGAAAGC 584

Qy 2308 AGAGACATTACTTGGCAACAAAGCCCATCTAGTCAAGGC---TATGTTTTTCCAGTG 2364
Db 593 AGAGATATTCTTTTACCAACAAAGCTCGCTAGTCAAGGCCTATGTTTTTCCAGTG 524

Qy 2365 GTATGATGATGTGAGATGTGACTGTGAAGAAAGCTGAGCACTGAAGATTTGATGCT 2424
Db 523 GTCATGTATGATGTGAGATGTGACTGTGAAG--AAGCTGAGCGCCAAAGAAATTTGATGCT 465

Qy 2425 TTTGAACCTGTGTGTTGGAGAGACTCTTGAGAGTCCCTTGGACTGCAAGGAGATCAAC 2484
Db 464 TTTGAACCTGTGTGTTGGAGAGACTCTTGATAGTCTTTTGGACTGCAAGGAGATCCAC 405

Qy 2485 CAGTCCATTTGAAGGAGATCAGCTCTGGGTGTTCTTTGGAAGGACTCATGCTGAAGCTG 2544
Db 404 CAGTCCATCCTAAAGGAGATCAGTCTGGGTGTTCTTTGGAAGGACTCATGCTGAAGCTG 345

Qy 2545 AAATCCAGTACTTTGGCCACTGTATCAGAGAGTGTGACTCACTGGAAAGACCTGTATG 2604
Db 344 AAGTCCCAATACCTTGGCCACTCATGCGAAGAGTTGACTCATTTGGAAGAAACCTGTG 285

Qy 2605 CTGGAGGAGATTGGGGCAGGAGAGAGAGGAGACAGAGAGATGAGATGCTGGATGGC 2664
Db 284 CTGGAGGAGATTGGGGCAGGAGAGAGAGGAGGAGAGAGATGAGATGCTGGATGGC 225

Qy 2665 ATCACTGATCGATGAGCTGATGCTGGGTGAATCTCCCTGGAGTTGGTATGACAGGGAG 2724
Db 224 ATCACTGATCGATGAGATGATGATTTGATGAACTCTGGGAGTTGGTATGACAGGGAG 165

Qy 2725 GCTGTCTGGCGGATTCATGGGTCAAAAAGTTGGACACGACTCAGCAACTGAACT 2784
Db 164 TCTGGCATGCTGCAATTCATGGGTGCGAAAGAGTTGGTACCACTGAGCGACTGAACT 105

Qy 2785 GAATGAACTGACTGAAACCTTAGTATGTTAT 2817
Db 104 GAATGAACTGAACTGAACTGATGTTCTCTTT 72

RESULT 3
US-08-756-506-23
; Sequence 23, Application US/08756506
; Patent No. 5905185
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Cottingham, Ian R.
; APPLICANT: Temperley, Simon M.
; APPLICANT: Foster, Donald C.
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Prunkard, Donna E.
; TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC
; ANIMALS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc. East
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,506
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 95-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6763 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-756-506-23

Query Match 5.4%; Score 194.3; DB 1; Length 6763;
Best Local Similarity 78.2%; Pred. No. 1.1;
Matches 279; Conservative 0; Mismatches 57; Indels 21; Gaps 4;

Qy 2453 TGAGAGTCCCTTGGACTGCAAGAGATCCAAACAGTCCATTTCTGAAGGAGATCAGCCCTG 2512
Db 5268 TGAGAACTGTCTGCACACAGCAGAGATCCACAGTC--ATCCTAAGGAGATCAGTCTG 5325

Qy 2513 GGATTTCTTTGGAAGGAATGATCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCA 2572
Db 5326 GTGTTCAAT--GGAGGACTGATGTTGAGCTGAAACTCCAACTGTTGGCCACCTGATG 5383

Qy 2573 GAAGAGCTGACTCACTGGAAAAGACCTGATGCTGGAGGGATTTGGGGCAGGAGAGAA 2632
Db 5384 GAAGAGCTGACTCATTTGAAAAGACCTGATGCTGGGAAAAGATTGAGGGCAGGAGAGAA 5443

Qy 2633 GGGGACACAGAGGATGAGATGGCTGGATGGCATCACTGATCGATGGAGCTGAGTCTGG 2692
Db 5444 GGGGACACAGAGATGAGATGGTGGATGGCATCAACACACANTGGACATGGGTTGG 5503

Qy 2693 GTGAACTCTCTGGAGTTGGTATGGACAGGAGCCCTGTCTGTC-----GGGATTCATG 2746
Db 5504 GTGAACTCTCAGGAGTTGGTATGGACAGGAGCCCTGGCTGTACCGAAGCGGTTTATG 5563

Qy 2747 GGGTCACAAAGAGTTGGACACAGACTGAGCAACTGAACTGAACTGAACTGACTGAAA 2803
Db 5564 GGGTCACAAA-----GACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5609

RESULT 4

US-08-206-176-7
; Sequence 7, Application US/08206176
; Patent No. 5639940
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Dalrymple, Michael A
; APPLICANT: Prunkard, Donna E
; APPLICANT: Foster, Donald C
; TITLE OF INVENTION: Production of Fibrinogen in Transgenic
; TITLE OF INVENTION: Animals
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,176
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: ovine beta-lactoglobulin

US-08-206-176-7
Query Match 5.4%; Score 194.3; DB 1; Length 10807;
Best Local Similarity 78.2%; Pred. No. 0.69;
Matches 279; Conservative 0; Mismatches 57; Indels 21; Gaps 4;
QY 2453 TGAGAGTCCCTTGGAGTGCAGGAGATCCAAACAGTCCATTCTGAAGGAGATCAGCCCTG 2512
Db 9312 TGAGAACTGTCTGCACACAGCAGAGATCCACCAAGTC--ATCCTAAGGAGATCAGTCTTG 9369
QY 2513 GGATTTCTTTGGAGGATGATGCTAAAGCTGAACTCCAGTACTTTGGCCACTGATCA 2572
Db 9370 GTGTTCAATT--GGAGGACTGATGTTGAAGCTGAACTCCAAATGCTTTGGCCACTGATG 9427
QY 2573 GAAGAGTCTGACTCACTGGAAGAACCTGATGCTGGAGGGGATTGGGGGAGGAGGAGAA 2632
Db 9428 GAACAGCTGACTCATTTTGAAGAACCTGATGCTGGGAAGATTTGGGACAGGAGGAGAA 9487
QY 2633 GGGACACACAGAGATGAGATGGCTGATGGCATCTGATGCTGATGAGAGTCTGG 2692
Db 9488 GGGACACACAGAGATGAGATGGCTGATGGCATCTGATGCTGATGAGAGTCTGG 9547
QY 2693 GTGAACCTCTGGAGTTGGTGTGATGACAGGAGGCTGCTCTGC-----GGCGATTTCATG 2746
Db 9548 GTGACTCCAGGAGTTGGTGTGATGACAGGAGGCTGCTCTGC-----GGCGATTTCATG 9607
QY 2747 GGGTCACAAAGATTGGACACGATGAGCAACTGAACTGAACTGAACTGAACTGAA 2803
Db 9608 GGGTCACAAA-----GACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 9653

RESULT 5

US-08-756-506-5
; Sequence 5, Application US/08756506
; Patent No. 5903185
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Cottingham, Ian R.
; APPLICANT: Temperley, Simon M.
; APPLICANT: Foster, Donald C.
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: Prunkard, Donna E.
; TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC
; TITLE OF INVENTION: ANIMALS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/756,506
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 95-28
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10807 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-756-506-5
Query Match 5.4%; Score 194.3; DB 1; Length 10807;
Best Local Similarity 78.2%; Pred. No. 0.69;
Matches 279; Conservative 0; Mismatches 57; Indels 21; Gaps 4;
QY 2453 TGAGAGTCCCTTGGAGTGCAGGAGATCCAAACAGTCCATTCTGAAGGAGATCAGCCCTG 2512
Db 9312 TGAGAACTGTCTGCACACAGCAGAGATCCACCAAGTC--ATCCTAAGGAGATCAGTCTTG 9369
QY 2513 GGATTTCTTTGGAGGAAATGATGCTAAAGCTGAACTCCAGTACTTTGGCCACTGATCA 2572
Db 9370 GTGTTCAATT--GGAGGACTGATGTTGAAGCTGAACTCCAAATGCTTTGGCCACTGATG 9427
QY 2573 GAAGAGTCTGACTCACTGGAAGAACCTGATGCTGGAGGGGATTGGGGGAGGAGGAGAA 2632
Db 9428 GAAGAGTCTGACTCATTTTGAAGAACCTGATGCTGGGAAGATTTGGGACAGGAGGAGAA 9487
QY 2633 GGGACACACAGAGATGAGATGGCTGATGGCATCTGATGCTGATGAGAGTCTGG 2692
Db 9488 GGGACACACAGAGATGAGATGGCTGATGGCATCTGATGCTGATGAGAGTCTGG 9547
QY 2693 GTGAACCTCTGGAGTTGGTGTGATGACAGGAGGCTGCTCTGC-----GGCGATTTCATG 2746
Db 9548 GTGACTCCAGGAGTTGGTGTGATGACAGGAGGCTGCTCTGC-----GGCGATTTCATG 9607
QY 2747 GGGTCACAAAGATTGGACACGATGAGCAACTGAACTGAACTGAACTGAACTGAA 2803

D_b 9608 GGGTCACAAA-----GACTGAGTGACTGAACTGAGCTGAATGGAAA 9653

RESULT 6

US-08-071-601-3
 / Sequence 3, Application US/08071601
 / Patent No. 5530177
 / GENERAL INFORMATION:
 / APPLICANT: BLECK, GREGORY T.
 / APPLICANT: BREMEL, ROBERT D.
 / TITLE OF INVENTION: DNA SEQUENCE ENCODING BOVINE
 / TITLE OF INVENTION: ALPHA-LACTALBUMIN AND METHODS OF USE
 / NUMBER OF SEQUENCES: 20
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: ANDRUS, SCARLES, STARKE & SAWALLI

ADDRESS: ANDRUS, WISCONSIN, STARKES & SAWALL
STREET: 100 E. WISCONSIN AVE., SUITE 1100
CITY: MILWAUKEE
STATE: WI
COUNTRY: USA
ZIP: 53202-4178
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/071.601

```

; APPLICATION NUMBER: US/08/071,601
;
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/744,765
; FILING DATE:
;

```

ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: F. 3262-1
TELECOMMUNICATION INFORMATION:

TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 255-2022
TELEFAX: (608) 255-2182
TELEX: 26832 ANDSTARK
INFORMATION FOR SEQ ID NO: 3:

```

; INFORMATION FOR SEQ ID NO: 3:
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 2044 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: signal peptide coding region
;

```

```

? NAME/ABB: signal peptide counting region
? LOCATION: 1943..2043
? FEATURE:
? NAME/KEY: inherited control region for a-lactalbumin
? LOCATION: 1966

```

```

1 LOCATION: 1960
2
3 FEATURE:
4
5 NAME/KEY: putative steroid response element
6
7 LOCATION: 1433..1446
8
9 FEATURE:

```

```

NAME/KEY: RNA polymerase binding region
LOCATION: 1961..1978
US-08-071-601-3

```

Query Match	5.1%;	Score 193.2;	DB 1;	Length
Best Local Similarity	81.1%;	Pred. No. 4.9;		
Matches 236;	Conservative	0;	Mismatches 53;	Indels

Qy 2502 GATCAGCCTGGATTCTTTTGAAGGAATGATCTAAAGCTGAAAT
Db 1 GATCAGTCTCTGGTGGTCATTGAAAGGACTGATGCTGAAGTTGAAGCTCT

Qy 2562 CCACCTGATCAGAGAGCTGACTCACTCGAAAGACCCCTGATGCTGTGCTGGG
Db 61 CCACCTGATGCGAAGAACTGACTCATGTGTAAGACCCCTGATCTGGG

Qy	2622	CAGGACGGAAGGGGACGACAGAGGATGACATCGCTGGATGGCATCACTCACTCGATGGA	2688
Db	121	CAGAGGAGAA-GGGATGACAGAGGATGGAGAGTTGGATGGATCAACCAACTCGATGGA	178
Qy	2682	CGTGA GTCTGGGTGAAC-TCTCTGGAGTTGGTGTATGACACAGGAGAGCCTGTCTCTGGCGGA	2740
Db	180	CATGAGTTTGAGCAAGCTTCACGAGGTGGTAAATGGGCAGGGAAGCCTGGCGTCTCGAG	239
Qy	2741	TTCAATGGGGTCAACAAGAGTTGGACACGACTGAGCAACTGAACCTGAACCTGA	2791
Db	240	TCCATGGGGTTGCAAGAGTTGGACACTACTGAGTGACTGAACCTGAACCTGA	290

RESULT 7

```

RESOLUT
US-08-621-100-3
; Sequence 3, Application US/08621100
; Patent No. 5850000
; GENERAL INFORMATION:
; APPLICANT: BLECK, GREGORY T.
; APPLICANT: BREMEL, ROBERT D.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING BOVINE
; TITLE OF INVENTION: ALPHA-LACTALBUMIN AND METHODS OF USE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

```

/ /
/ ADDRESS: ANDRUS, SCALES, STARK & SAWALL
/ STREET: 100 E. WISCONSIN AVE., SUITE 1100
/ CITY: MILWAUKEE
/ STATE: WI

STATE: WI
COUNTRY: USA
ZIP: 53202-4178
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

?
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
?

```

```

? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/621,100
? FILING DATE: 22-MAR-1996
? CLASSIFICATION: 800
? PRIOR APPLICATION DATA:
?

```

```

? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/071,601
? FILING DATE:
? APPLICATION NUMBER: US/07/744,765
? FILING DATE:
?

```

ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: F. 3262-1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 255-2022
TELEFAX: (608) 255-2182
TELEX: 26832 ANDSTARK

```

, INFORMATION FOR SEQ ID NO: 3 :
,-----
, SEQUENCE CHARACTERISTICS:
,-----
, LENGTH: 2044 base pairs
, TYPE: nucleic acid
, STRANDEDNESS: single

```

```

: STRANDEDNESS: single
:
: TOPOLOGY: linear
:
: MOLECULE TYPE: DNA (genomic)
:
: FEATURE:
: NAME/REV. signal peptide coding region

```

```

; NAME/KEY: signal peptide coding region
; LOCATION: 1943..2043
;
; FEATURE:
; NAME/KEY: inherited control region for a-lactalbumin
; LOCATION: 1966

```

```

; LOCATION: 1966
;
; FEATURE: -
; NAME/KEY: putative steroid response element
; LOCATION: 1433..1446
;
; FEATURE:

```

```

; FEATURE:
; NAME/KEY: RNA polymerase binding region
; LOCATION: 1961..1978
US-08-621-100-3

```

Query Match	5.1k; Score 183.2; DB 1; Length 2044;
Best Local Similarity	81.1k; Pred. No. 4.9;
Matches 236; Conservative 0; Mismatches 53; Indels 2; Gaps 2;	
Qy	2502 GATCAGCCCTGGGATTTCTTTTGGGAAGAAATGATCTAAAGCTGAAACTCCAGTACTTTGG 2561
Db	1 GATCAGTCTCTGGGTGTCATTGAAAGGACTGATGCTGAAGTGAAGCTCCAATACTTTGG 60
Qy	2562 CCACCTCATCAGAGAGCTGACTCACTCGGAAGAGACCCTGATCTGGAGGGATTCGGGG 2621
Db	61 CCACCTGATGCGAAGAACTGACTCATGTGATAGACCCTGATCTGGGAAGATTGAAGG 120
Qy	2622 CAGGAGGAGAGGGGACGACAGAGGATGAGATGGCTGGATGGCATCACTCACTCGATGGA 2681
Db	121 CAGGAGGAGAA - GGGATGCAGAGGATGGAAGAGTTGGATGGAATCAACCACTCGATGGA 179
Qy	2682 CGTGAGTCTGGGTGAAC - TCTGGAGTGTGTGATGGACAGGGAGCCCTGCTTCGCGCGA 2740
Db	180 CATGAGTTTGAGCAAGCTTCAGAGTTCGTAATGGCAGGGAAGCTTCGCGTCTGCAG 239
Qy	2741 TTCATGGGGTCAAAAGAGTTGGACACGACTGAGCAACTGAACTGAAGTGA 2791
Db	240 TCCATGGGGTTCAAAGAGTTGGACACTCTGATGTCGACTGAACTGAAGTGA 290

```

RESULT 8
US-09-930-377B-1
; Sequence 1, Application US/09930377B
; Patent No. 6677500
; GENERAL INFORMATION:
; APPLICANT: Wheeler, Matthew B.
; APPLICANT: Donovan, Sharon M.
; APPLICANT: Black, Gregory T.
; APPLICANT: Monaco-Sigal, Marcia
; TITLE OF INVENTION: Animals Expressing Exogenous IGF-I in their Milk
; FILE REFERENCE: 66-00
; CURRENT APPLICATION NUMBER: US/09/930,377B
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,474
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4532
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:alpha-LA/IGF-I
; OTHER INFORMATION: gene construct
US-09-930-377B-1

```

Db 240 TCCATGGGGTTGCAAGAGTTGGACACTACTAGTGACTGAACCTGA 290

RESULT 9
US-09-193-562D-1/C
; Sequence 1, Application US/09193562D
; Patent No. 6309857
; GENERAL INFORMATION:
; APPLICANT: Pauli, Benedicht U.
; TITLE OF INVENTION: Nucleotide Sequences Encoding Mammalian Calcium
; TITLE OF INVENTION: Activated Chloride Channel-Adhesion Molecules
; FILE REFERENCE: 18617.0052
; CURRENT APPLICATION NUMBER: US/09/193,562D
; PRIOR FILING DATE: 1998-11-17
; CURRENT APPLICATION NUMBER: US/60/065,922
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 3317
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: sequence encoding Lu-ECAM-1 and Lu-ECAM-1 associated
; OTHER INFORMATION: protein from bovine endothelial cells
US-09-193-562D-1

RESULT 10
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 593827
; GENERAL INFORMATION:
; APPLICANT: Sam, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun

QY 3080 AGTGTTCACCAAGTCTTTAGGCTACTCTTTTGATTTATTTGGTCTATCTTTCTCA 3139
DB |||||
QY 4597 TGTGTGGGTAGAAATATATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4656
DB |||||
QY 3140 AGTTTGAATTTGGCTACCACTCAATTTATCTTTATTTTGTAAATAGCTCTTTAAAT 3199
DB |||||
QY 4657 TTAATTTAAGGTGTTTGTGGAATGATTTATTTATTTTAAATTTGATTTAAAT 4716
DB |||||
QY 3200 TCATTAATCTTTGATAACAGCTTCAGTTCATAGGCTTTTAAATAAGTCTTTTGT 3259
DB |||||
QY 4717 TTTTAAATTTTGTATTTATGTTAGATATGTTAAACGAATGAAGTCTTTTGAATAAT 4776
DB |||||
QY 3260 TTTTAAAGAAATGTCATCTTTCTGGAAGTTTTCACATGCTTTGACCAATAATTTAGGA 3319
DB |||||
QY 4777 TTTGT--ATGATTTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 4834
DB |||||
QY 3320 TATTTTGAATGGTTCATAGTATGCTTTTGTACTTGGCAATTTATGGAAGTTTATGATTT 3379
DB |||||
QY 4835 TATGTTTCATTTTGAATTTTATTTAGTAAGTATAGTTAGTATTTTAAATGTTCTTA 4894
DB |||||
QY 3390 ATGAATATGATGCTTTTGTGGCATAAAGGCTATGCAATTTTGTGCTCTATA 3439
DB |||||
QY 4895 TAGTATGCTATTTTGTGTATAATTTTATATGTTTATTTAGTATGTTGTTTATTTT 4954
DB |||||
QY 3440 TTTCTAAATTTATA 3453
DB |||||
QY 4955 TTTGATAAATAATA 4968
DB |||||

RESULT 12
US-10-204-708-63
; Sequence 63, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEBENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 63
; LENGTH: 5562
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-63

Query Match 2.8%; Score 100.9; DB 1; Length 5562;
Best Local Similarity 45.1%; Pred.No.21;
Matches 342; Conservative 0; Mismatches 336; Indels 19; Gaps 3;
QY 2789 TGAATCTACTGAACCTTAGTACTTTATATCTACAGAAATAGTAATTTCTATATGTAT 2848
DB |||||
QY 1239 TGTAGTAGTAGAGATCGCGTTATTTGTTATTTTATTTTGGGTAGTAGCGAGATTGT 1298
DB |||||
QY 2849 TCAAAATTTTATATATGTTGGTTAGATAAATAGATTTTCAATTTGATTTTATCTTT 2908
DB |||||
QY 1299 TTAATAAGAAAAAATGATTTCTTTATATTTTAAAGTTATATTTATTTTGTGTTG 1358
DB |||||

QY 2909 GATTTTCTCTACTTATTTAAATTTTGGGATTTTAACTATTTCTTCAATGACTTGTATTC 2968
DB |||||
QY 1359 GGGTGAATTTTGT 1418
DB |||||
QY 2969 TAATATTTACTTATCTATTTTACCTTAAATTCACCTTATTTTATGATTTTCTTAATAA 3028
DB |||||
QY 1419 CGGTTAGGAATGATTTTGTGATTTATTTATTTTATTTTATTTTATTTTATTTAGTTAGT 1474
DB |||||
QY 3029 AATCCAGTCCCTTGT 3086
DB |||||
QY 1475 -----TT 1524
DB |||||
QY 3089 CCAGTCTCTTCAGGCTACTCTCTTTTGATTTATTTTGGTCTCTCTTTCTCAAGTTTTCGA 3148
DB |||||
QY 1525 TTTTGT 1584
DB |||||
QY 3149 TTGGCTACGTAACTCATTTATCTTTTATTTTGTAAATTAGCTCTTTAAATTCATTTATTC 3208
DB |||||
QY 1585 TAAGAAAGTTTATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAGGATAGTGTAT 1644
DB |||||
QY 3209 TTTGATAACAGCTTCAGTCTCATGCTTTTAAAGTTTATTTTATTTTATTTTATTTTAA 3268
DB |||||
QY 1645 TTTTGAATTTTGTGATATAAAGATTTTATTTATTTATTTATTTATTTTATTTTATTTTATTT 1704
DB |||||
QY 3269 GAATGTCTCTTTCTGGAAGTTTTCACAAATGCTTTGACAAATAATTTAGGATATTTTGA 3328
DB |||||
QY 1705 TTATGATGCTTTTGT 1760
DB |||||
QY 3329 ATGTTTCATGATATGCTTTTGTACTTGGCATTTTGAAGTTTATGATTTATGATTTAT 3388
DB |||||
QY 1761 TTTGAAATTTATGATGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTGAAGGTATAT 1820
DB |||||
QY 3389 GATGCTTTTGTGGCATAAAGGCTTA--TGSCATATTTTGTGTCTATTTCTTCTTAA 3447
DB |||||
QY 1821 TTATTTTATTTAGTTTGTGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1880
DB |||||
QY 3448 ATTATAAATGCTTTTAAAGATTTTCTGCTATTA 3484
DB |||||
QY 1881 AGTATATATGCTTTTGT 1917
DB |||||

RESULT 13
US-10-204-708-10
; Sequence 10, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEBENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 10
; LENGTH: 6070
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-10

```

; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 80
; LENGTH: 8961
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (5096)
; OTHER INFORMATION: n is a or g or c or t
US-10-204-708-80

Query Match          2.8%;   Score 100.1;   DB 1;   Length 8961;
Best Local Similarity 47.9%;   Pred. No. 14;   Indels 7;   Gaps 2;
Matches 345;   Conservative 0;   Mismatches 369;

QY 2786 AACTGAACGTACTGAACCTTAGTAGTTATATATTAATCTACGAAAAATAGTAATTCATATG 2845
Db 5327 AATAGATATTTTGTGTAATTTTTTTTTTATTTTGGATTTAATGTTAATTTTATAAT 5386

QY 2846 TATTCAAAATATTTTCATAATGTTGGCTTAAGATAATAAGATTTTCAAATTGATTTTATC 2905
Db 5387 AATTTATATTTTTTTTTTTTAAGTGTATTTTATTAATAGTATTTTTCGTAGTATTTAT 5446

QY 2906 TTTGATTTTTCTCTACTATTTAAATTTGGGATTTTAACTATTTCTTCAATGACTGTAT 2965
Db 5447 GTTTTATATTTTATAATTTAGTTTTATTTTTATAGATATTTTATAAATATTTT 5506

QY 2966 TTCTAATATTTACTTATCTATTTTACTTTAATTCGACTATTTTATTTGATTTTCTAA 3025
Db 5507 TATTTATTTTAGTTTGTGATTTTTATAAGTGTAGAAATTTATTTTTTTTATATATAA 5566

QY 3026 TAAATCCAGTCCCTGTTTTTTTAAAAAGACTTTAAAAATATTA-ATTTCTCTTTAGTGT 3084
Db 5567 TAACGGATATTCGGTTTTTTTTTATATATATTTTCGTAATAATTAGTTTTTATTTTAT 5626

QY 3085 TTTACAGGTTCTTTCAGGCTACTCTTTTGATTTATTTGGTCCATCTTTCTCAAGTTT 3144
Db 5627 ATTGATATTTTATAAATAATTAATTTTTTAAAGTGTGTTTTTAGTTTGTAGGTAGGGTT 5686

QY 3145 TGAATTTGGCTACGTAACATTTATCTTTATTTTTTGTAAATAGTCTTTAAATTCATT 3204
Db 5687 TAGAGGTTATATATTTTAGTTTTTTTTTTTTTTTAGTTTATAGTTATAGTAAGTATT 5746

QY 3205 ATTCTTTGATTAACAGCTTCAGTTCTAAGGCTTTAAATAAGGTTTTTTTTTTTTTTTT 3264
Db 5747 TTTTTTTTATATAAAAAGATAAATTATAAATGATTAATAGTTTATATGTCGTTTTATAGATG 5806

```

[illegible]

```
QY 3505 T 3505
Db 6041 T 6041

RESULT 15
US-10-204-708-29
; Sequence 29, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 29
; LENGTH: 5666
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-29

Query Match 2.7%; Score 97.1003; DB 1; Length 5666;
Best Local Similarity 51.5%; Pred. No. 23;
Matches 333; Conservative 0; Mismatches 294; Indels 19; Gaps 5;

QY 2842 TATGTTTCAAAATTTTCTCATATGTTGGTTAGATAAATAGATTTTCAAAATGATTTT 2901
Db 82 TGTGATTTTGAAGAATTTATTTATTTTATATAATTTATAGGTTTANGAGGTAAAGTT 141

QY 2902 TATCTTTGATTTTCTCTACTATTATTAATTTTGGGATTTTAACTATTCTTCAATGACCT 2961
Db 142 GGTTTTAGTGTGTTTTTTTATTAATTTT-TTTTAAATGTTCCGATGTTTGTGGTATTT 199

QY 2962 GTATTTCTAATTTACTTATCTATTTTACTTTTAACTGACATTTTATTTGATTTT 3021
Db 200 TTTATTTTATTAATGATGATGTTGTTGTTATTTT-TTTTATTTTATTTAGTTT 259

QY 3022 CTAATAAAATCCAGCTCTGTTTTTTTAA-----AAAAGACTTTTAAATTTAAATTTCT 3075
Db 260 TTAGATTTGTTGATTTTATTTGATTTT-TTTTAAAGATAAATTTTGTTTTATTTGATTTT 319

QY 3076 CTTTAGTGTTTTACAGCTCTCTTCAGGCTACTCTCTTTTGTATTTTATTTGGTCTATCTTT 3135
Db 320 TTTATGCTTTTTTATTTTAAATTTTATTTGATTTTGTGTTTAAAGTTTATTTGTTTTTTT 379

QY 3136 CTCAAGTTTGAATTTGGCTAGTAACATTTATCTTTATTTTGTGTAATTTAGCTCTTT 3195
Db 380 TTTTGTGTTTTTGGGTTAGTTGTTTATTTTGTATAGATTTTGTGAAGTATAATTTA 439

QY 3196 AAATCAATATTTCTTGATACAGCTTCAGTCTCTATGCTTTTAAATTAAGTTTTTTT----- 3250
Db 440 GAATATTGAATTTGAAGTTTGTGTTTTTAAATTTTAAATATGGGTATAAATTTT-TTTTAT 499

QY 3251 - - - - -TTTTTTTAAAGATGTCATCTTTTGTGAAGTTTGTGAAGTCTTTTGTGCA 3308
Db 500 TGTTTTAGTTGATTTTATAAAATTTTGATGAGGATTTTATTTTATTTAGATTTGAGAT 559

RESULT 16
US-10-204-708-20
; Sequence 20, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 20
; LENGTH: 6866
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-20

Query Match 2.7%; Score 96.1003; DB 1; Length 6866;
Best Local Similarity 47.4%; Pred. No. 20;
Matches 317; Conservative 0; Mismatches 349; Indels 3; Gaps 1;

QY 2813 TTTATTTACTCAGAAATAGTAAATTTTCAATGATTTTCAAAATTTTCAATAAATGTTGGT 2872
Db 566 TTGTTTTTAAAGAGATGTTGGTTATTTTATTTAGTGATATTTTTCGGTTATTGA 625

QY 2873 TAAGATAATAAGATTTTCAAAATTTGATTTT-TTTATCTTTGATTTTCTCTACTTTATTA 2929
Db 626 TATATTTAGGGGATTTATAGATTTGTTGCGATTAAGTTGTTTTTTTAAATATTTTAA 685

QY 2930 TTTTGGGATTTTAACTATTCTTCAAGACTTGATTTTCTAAATTTTACTTATTTCTATT 2989
Db 686 GATTTATGCTAGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 745

QY 2990 TACTTTAATTTGACATTTTATTTTATTTGATTTTCTTCAATAAATCCAGCTCTGTTTTTTA 3049
Db 746 AATTTTGTTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 805

QY 3050 AAAAGACTTTTAAATTTAATTTTCTCTTTAGTGTTTTACAGGTTCTTTCAGGCTACTTC 3109
Db 806 AATATAGTTTGGTGAATGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 865

QY 3110 TTTTGTATTTATTTGTCCTACTCTTTTCTCAAGTTTTCGAATTTTGAATTTGGCTAGTAACT 3169
Db 866 TAGGTTTTAGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTATTT 925
```


QY 2866 TGTGTTAAGATAAAGATTTTCAAATTCATTTTATCTTTGATTTTCTCTACTTAT 2925
 Db 5266 TTAAGTTTAAGTAGGAGAAAGTTTATAAAATTTTATTTATGAATTTTGTATATC 5325
 QY 2926 TTAATTTGGGATTTAACTATTTCTTCAATGACTGTTGATTTCTTAATTTACTATTCT 2985
 Db 5326 GTATTTTAAAGGATTTTATTTTAAATGTTTAGAGTTTGTGTTTATATATAA 5385
 QY 2986 ATTTTACTTTAATGCACTTTATTTATGATTTTCTTAATAATCCAGTCCTTGT 3045
 Db 5386 TAGTTCGTTATTAAGATTTTATTTTANGATGTTGTAGATAAGTTTAAATTTTGT 5445
 QY 3046 TTTAAAAAGACTTTAAATTTAATTTCTCTTTAGTGTTTTACCAGTTCTTTCAGGCTA 3105
 Db 5446 TTTGTTTAAATTTAGTTGTTTAAATGATTTGATTTGATTTATTTTGTATTTTA 5505
 QY 3106 CTTCTT-----TTGATTTATTTGCTCTATCTTTCTCAAGTTTGAATGGC 3153
 Db 5506 TTTGTTATTTAAGTTTATAGATGATGTTGGAGTTAATTTTATTTTATGATTTATTTT 5565
 QY 3154 TACGTAACCTATTTATCTTTATTTTGTAAATAGCTCTTTAAATTCATTTCTTTGA 3213
 Db 5566 TTAATTTTGAATTTATTTATTTTAAATAGTAGTTTATTAATTTTGTAT--G 5623
 QY 3214 TAAAGCTTCAGTTCTATGCTTTTAAAGTTTCTTTGAGCAATAATTTAGATTTTGAATGGT 3273
 Db 5624 TTTGGGGTGTGTTTTCGGGAATGTTTGGTATTTTGTGTTTGTGTTTGTGTTTGTATG 5683
 QY 3274 TCATTTCTTGTGAAGTTTGTGCAATGCTTTGAGCAATAATTTAGATTTTGAATGGT 3333
 Db 5684 AATGAGTTTATGATTTTATTTTAAATAGTAGTTTATTAATTTTGTATTTTGTATG 5743
 QY 3334 TCATGAGTATGCTTTGCTGCTGCTTTTATGAGTTTATGATTTTATGATTTATGATG 3393
 Db 5744 ATTTTGTGATGTTTATTTAGGAAAGTAAATTTGAATGTTAAATTTTATTTTGTAGT 5803
 QY 3394 TTTTTTTGGCATAAAGGCTATGCGATATTTTGTGCTATATTTCTTAAATATA 3453
 Db 5804 TATTTTGGTTTTAAATTTTAAAGGTTTATATTTTATATAAAGTTTGT 5863
 QY 3454 AATTTGGCTTTAAAGATTTTGTGCTATT 3483
 Db 5864 ATTTTATGTTTATGATTTTATTTTATGATTT 5893

RESULT 19

US-08-998-416-595
 ; Sequence 595, Application US/08998416
 ; Patent No. 6239264
 ; GENERAL INFORMATION:
 ; APPLICANT: Philippsen, Peter
 ; APPLICANT: Pohlmann, Rainer
 ; APPLICANT: Steiner, Sabine
 ; APPLICANT: Mohr, Christine
 ; APPLICANT: Wendland, Jurgen
 ; APPLICANT: Knechtle, Philipp
 ; APPLICANT: Reibschung, Corinne
 ; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
 ; TITLE OF INVENTION: AND USES THEREOF
 ; NUMBER OF SEQUENCES: 1152
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: No. 6239264artis Corporation
 ; STREET: 3054 Cornwallis Road
 ; CITY: Research Triangle Park
 ; STATE: No. 6239264th Carolina
 ; COUNTRY: USA
 ; ZIP: 27709
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/998,416
 ; FILING DATE: 24-DEC-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: CH 0016/97
 ; FILING DATE: 31-DEC-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meigs, J. Timothy
 ; REGISTRATION NUMBER: 38,241
 ; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 919-541-8587
 ; TELEFAX: 919-541-8689
 ; INFORMATION FOR SEQ ID NO: 595:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 658 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; ORIGINAL SOURCE:
 ; ORGANISM: PAG1408RP
 ; US-08-998-416-595

Query Match 2.7%; Score 94.7002; DB 1; Length 658;
 Best Local Similarity 50.9%; Pred. No. 2e+02;
 Matches 248; Conservative 0; Mismatches 238; Indels 1; Gaps 1;

QY 2810 TAGTTTATATTAATTAAGATTTTCAAATTCATTTTATCTTTGATTTTCTCTACTTATTTA 2928
 Db 171 TTGTATTAATTAACCATTAATTAATTTATCTTGAATTAATTTATTTTGTATTTATTT 230
 QY 2870 -GGTTAAGATAAATTAAGATTTTCAAATTCATTTTATCTTTGATTTTCTCTACTTATTTA 2928
 Db 231 ATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 290
 QY 2929 ATTTTGGGATTTTAACTATTTCTTCAATGACTTGTATTTCTTAATTTACTTATTTCTATT 2988
 Db 291 ATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 350
 QY 2989 TTACTTTAATTCGACTTATTTTATTTATTTGATTTTCTTAATAAATCCAGTCCTTGTGTTTTT 3048
 Db 351 TTTATTTATTTGTTGTTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 410
 QY 3049 AAAAAGACTTTAAATTAATTAATTTCTCTTTAGTGTTTTACAGTCTCTTTCCAGGCTACTT 3108
 Db 411 TTTTATTTTATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 470
 QY 3109 CTTTGTGATTTTATTTGGTCTATCTTTTCTCAAGTTTGTGAATTTGGCTAGTAACCTCATTTA 3168
 Db 471 ATTTGTTGACTTTATTTCTTATTAATAGATTTGATTTTGTGAACATATATATATATGCTATT 530
 QY 3169 TCTTTATTTTGTAAATTAAGTCTTTTAAATTCATTTCTTTGATTAACAGCTTCAGTTC 3228
 Db 531 TCTATTTCTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 590
 QY 3229 TATGCTTTAATAAGTTTATTTTATTTTATTTTATTTTATTTTAAAGATGTCATTTCTTTGTGAG 3288
 Db 591 TTTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 650
 QY 3289 TTTTGAC 3295
 Db 651 TTTTTCAC 657

RESULT 20

US-10-204-708-22
 ; Sequence 22, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENROCK, Christian
 ; APPLICANT: BERLIN, Kurt

RESULT 21
US-10-204-708-57
; Sequence 57, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 22
; LENGTH: 11049
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-22

Query Match 2.6%; Score 93.9003; DB 1; Length 11049;
Best Local Similarity 49.1%; Pred. No. 13;
Matches 300; Conservative 0; Mismatches 306; Indels 5; Gaps 2;
2858 TTTCATTAATGTTGGTTAAGATAAAGATTTTCAAATGATTTTATCTTTGATTTTCTTAAATTTA 2917
1191 TTGTGTAATATGTTAAAGATTTATTTTATTAGTAGTATTTTATTTAGGTTTATTTTATTT 1250
2918 CTACTTATTTAAATTTTGGGATTTTAACTATTTCTCAATGACCTGTTATTTCTAATTTA 2977
1251 TTTTGTGTAATTTAAATTTTATTTTATTTTATATATATTTACGTTTGTATTTATGT 1310
2978 CTTATTCATTTTACTTTAAATGTCACATTTATTTTATGATTTTCTTAAATAAAATCCAGTC 3037
1311 ATAATGTTGGTTAGACGACGTTTTCGATTTTAAATAGAGAGTTTAGATGACGTTTAA 1370
3038 CTGCTTTTATTAAGACATTTAAATATTTATTTCTTCTTTAGTGTGTTTACCAGTTCTT 3097
1371 TTTTATATTTATAGATATTTTAAATTTAAATGATTAATGAAATTTATGTTTATTTT 1430
3098 TCAGGCTACTTCTTTGATTTATTTGGTCTATC-TTTTCTCAAGTTTGAATGGCTAC 3156
1431 TATTTAGTTGTTATGTTTATTTTAGTGAATGTTTGTATTTATTTAGTTAGTAA 1490
3157 GTAACTCATTTATCTTTATTTTGTAAATAGCTCTTTAAATCATTTATCTTGATAA 3216
1491 GTTGGGTTATTTATTTATAGTTATTTAGGTTTATTAATTTTATTTTATTTTATTT 1550
3217 CAGCTTCAGTTCTATGGCTTTAAATAAGTTTATTTTATTTTATTTTATTTTAAAGAAATGCA 3276
1551 TGAAGGTATATATAGATTTTATTAACGTATTTTGGTGAAGATTTTAAAGTTTATTA 1610
3277 TTCTTTGGAAGTTTGTCAATGCTTTGAGCAATTAATTTAGGATATTTTGAATGGTTCA 3336
1611 TTGTTTTAAGATATTAATTTAAATTTTGTGTTAGGAAGTTTGAATGATTTG 1670
3337 ---TGAGTATGCTTTTGTACCTGGCATTTATGAGTTTATGATTTATGATTTATGATG 3392
1671 GTTTTGTGTTATTTTGTGTTATGTTTATTTTAAATTTAGAAATTTATTAAGTTTATTT 1730
3393 CTTTTTTTGGGCATAAAGGCTATGGCATATTTTGTGCTATATTTCTTAAATATAT 3452
1731 TTGTTTTAGGTTTATTTTATATGTTTATTTTGTGTTTGAATAGTTTATTTTATTTT 1790
3453 AAATGGCTTT 3463
1791 CGTTTAGTTT 1801

Query Match 2.6%; Score 93.8003; DB 1; Length 6326;
Best Local Similarity 50.4%; Pred. No. 23;
Matches 253; Conservative 0; Mismatches 247; Indels 2; Gaps 1;
2885 ATTTTCAAATTCATTTTATCTTTGATTTTCTTACTTATTTAAATTTTGGGATTTTAA 2944
2523 ATGTTTAGTTTGTGTTAGGTTATTTTGAATTTATTTTATTTTATTTTATTTAGTTG 2582
2945 TATTTCTTCAATGACCTGTTATTTCTAATATTTTACTTATTTTACTTTTAAATGCACT 3004
2583 TTATTTAGTTTGTGTTAGTTTGTATTTTTCGATTTTATTTTATTTATTTAGTTAGG 2642
3005 TATTTTATGATTTTCTTAAATAAATCCAGTCCTTGTGTTTTTAAAGACATTTAAAT 3064
2643 TTATTTATTTTATTTTATGTTAGTGTAGTACAGAGGTTTATTTAAATTTTGAAGTAGAT 2702
3065 TATTAATTTCTCTTTAGTGTGTTTACCAGTTCTTTCCAGGCTACTTCTTTTGAATTTATTTG 3124
2703 ATTTATATTTATGTAATTTTGTGTTTTTAGTATTTTACGATAGATTTATTTT 2762
3125 TCCTATCTTTTCTCAAGTTTGAATTTGGCTACGTAACCTCATTTATCTTTTATTTTGTAA 3184
2763 TATATTTTGTGTTTTTTTATAGTTTATGAGTTTTTTTGTGTTTATTTTATTTTATTTTAA 2822
3185 ATTAGCTCTTAAATTCATATTTCTTTGATAACAGCTTCAGTTCTATGCTTTAATAAAG 3244
2823 GGATTTTATTTTATTTGTTTTTGTGTTTGGATTTTTTTTGTAGGTGTTATTTAGTTTG 2882
3245 TTTTTTTTTTTTTTTTTTTTTTAAAGAAATGCTATTTCTTGTGAAGTTTTCGCAATGCTTTG 3304
2883 TTTTGTGTTTTTTTGTGTTTTATATATATAGTATGAGTTTTTGGAGGGTAGGATTTGGTGT 2942
3305 AGCAATAATTTAGGATATTTTGAATGGTTCAAGATATGCTTTTGTCTTGTGCAATTTAT 3364
2943 TGTTTTGTGTTT--GTATTTTGTATTTTGGATTAGGTTAAGTATGTTAGGTATTTAA 3000
3365 TGAAGTTTATGATTTATGAAT 3386
3001 TAAATAATGGATTTATAGTT 3022

US-10-204-708-57
; Sequence 57, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 57
; LENGTH: 6326
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-57

```
RESULT 22
US-10-204-708-2
; Sequence 2, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 2
; LENGTH: 10467
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-2

Query Match      2.6%; Score 92.7003; DB 1; Length 10467;
Best Local Similarity 49.1%; Pred. No. 15;
Matches 323; Conservative 0; Mismatches 328; Indels 7; Gaps 3;

QY 2807 TAGTAGTTTATATTACTCAGAAAATAGTAATTTTCATATGATTCACAAATATTTCATAAT 2866
DB 1503 TTGTAGTATAAAATAGTTTATATATATATGTAATAATTTTATTAATTTATTTATAT 1562

QY 2867 GTTGGTTAAGATAATAGATTTTCAAAATCGAATTTTAACTTATCTGATTTTCTACTAT 2926
DB 1563 GAAGTTTATTTATATAAAATAGTAGTTTGTATTTGTGGGTATAGTTTTCGATTTTG 1622

QY 2927 TAATTTTGGGATTTTAACTATTTTC---TTCAATGACTTCTATTTCTAAATTTACTAT 2982
DB 1623 TTTTATTTTGAATGATTTTGTGTTTATTAATTTATTTTATTTTATTTTGTGGGAAT 1682

QY 2983 TCTATTTTACTTTAATGCACTTATTTTATTTATTTTCTAATAAAATCCAGTCTCTGT 3042
DB 1683 TTGTTTTCGTATGATTTTCGGTGAAGTTGTAGTAAGTTATTTTGTATTTTATTTT 1742

QY 3043 TTTTAAAGAACTTTTAAATTA-TTAAATTTCTTTAGTGTTTTACAGTTCCTTTCAG 3101
DB 1743 GATTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAA 1802

QY 3102 GCATCTTTCTTTTATTTTGTCTTATCTTCTCAAGTTTGTAAATGGCTACCTAAC 3161
DB 1803 GTTATATGTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAGTTTGTAGTGTGGAAGATT 1862

QY 3162 TCATTTATCTTTATTTTGTGTAATAGTCTTTT-AAATCTATTTCTTTGATNAACAG 3219
DB 1863 TTGTGTTAATAGGGTTTATAAAATTTTTTTTTTTTGAATTTTATTTGGTTGTG 1922

QY 3220 CTTTCAGTTCTATGGCTTTAATAAGTTTTTTTTTTTTTTTTTTTTTTTTTAAAGATGCTATC 3279
DB 1923 TGGTGTAAATCTGTTTAAATAATATGTTTATTTAGTTTATTTGGATTTAGTATTTT 1982

QY 3280 TTGTGAAGTTTGAATGCTTTTGAACAAATATTAGGATATTTTGAATGGTTCAGTA 3339
DB 1983 TTAATAAAATTTTTTTTTTAGGGTTTGGTGATATAGTAAAGGTTTTTATTTTTTTTTT 2042

QY 3340 GTATGCTTTGTACTGGCATTTATTTGAAGTTTATGATTTATGATTTATGATGCTTTT 3399
```

```

2043 TTTTGTGTTTTTTTTTTTTTTTATTTTATATTTTAAATAGTTATTGTTATACGGATTTA 2102
QY 3400 TTGGGCATAAAGCTCATGGCATATTTTTTTGGTCTATATCTTAAATATATAAATT 3457
DB 2103 TTTTATATATAATAGTTTATTTAAGGTTTTTAAGGATTTTTTTTTTTTATATATT 2160

RESULT 23
US-10-204-708-69
; Sequence 69, Application US/10204708
; Patent No. 667731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 69
; LENGTH: 6040
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-69

Query Match      2.6%; Score 92.4003; DB 1; Length 6040;
Best Local Similarity 50.3%; Pred. No. 25;
Matches 320; Conservative 0; Mismatches 296; Indels 20; Gaps 4;

QY 2806 TTAGTAGTTTATATTACTCAGAAAATAGTAATTTTCATATGATTCACAAATATTTCATA 2865
DB 1189 TTAATAAATTTATTTAAGTTGGAGTTGTGTTTATTTATTTTATTTAAGTATTTTAAAC 1248

QY 2866 TGTGGTTAAGATAATAAGATTTTCAAAATGATTTTATCTTTGATTT---TTCTCTACT 2922
DB 1249 GATTTTTTTAGTTTTTTTTTTTTTTTTTTTAAATTTGTTTATTTAGAGTTGAGATTTAAT 1308

QY 2923 TATTTAATTTGGGATTTTAACTATTTCTTCAATGACTTGTATTTCTAATATTACTAT 2982
DB 1309 TATTGTATATTGAATATATAGTATTTGTTTAAAGTTGGTTTTTTTAGATTTTAAATTTT 1368

QY 2983 TCTATTTTACTTTTATGCACTTATTTTATTTGATTTTCTTAAATAAATCCAGTCTTGT 3042
DB 1369 TTTATTTAAATATATTTTATGTTTATTTAGTGGTTTTTTTAAATGATATATTTTATT 1428

QY 3043 TTTTAAAAAGACTTTAAAAATTTAATTT-----TCTCTTTAGTGTTTTACCAGTTC 3095
DB 1429 TTTTGTGTTTTAAAGTTTTTATTAAGTAAGTAATTTAAATATTTTAAATAATAAAGTTTTT 1488

QY 3096 TTTCAGGCTACTTCTTTGATTTATTTGGTCCCTATCTTTCTCAAGTTTGAATGGCTA 3155
DB 1489 GGTATGGTATAAAGTTTTTTAAAAATTTGGTTTTTAAATTTTATTTAGTTTATTTAT 1548

QY 3156 CGTAACCTCATTTATCTTTATTTTGTAAATTTAGTCTTTTAAATTTCAATTTCTTTGATA 3215
DB 1549 TTTATATATTTGATTTTTTTTTTTTATTTTCTTATATTAAGTTTTTATGTAATTATTG--- 1605

QY 3216 ACAGTTCAGTCTCATGGCTTTAAATAAAGTTTTTTTTTTTTTTTTTTTTTAAAGAATGTC 3275
```

Db 1606 --AAATTTATGTTTTTTTATATAAATGTTATTTTTTTTATTTTACGGATAGAGATTG 1663
QY 3276 ATCTTTTGTGAAGTTTTTGACAATGCTTTGAGCAATAATTTAGGATATTTTGAATGGTTC 3335
Db 1664 TTGTTTATTTTATTTAGTTAATGATTTTTTTTTTTTAAATATTTTTTAAAGT-----ATAGGTGA 1718
QY 3336 ATGAGTATGCTTTTGTACTTGGCAATTTATGAAGTTTATGATTTATGAATTTATGAGCTT 3395
Db 1719 GAGAGAAATATTTTTTTTTTTGGGATATTTTTTAAAGTAGTTTTTTAGTTAGTATTATTTT 1778
QY 3396 TTTTGTGGCAATAAAGTCTATGGCAATTTTTTTTGT 3431
Db 1779 TTTATTTGTGTTAATGATTTTTTATATATTTTTTTTGT 1814

RESULT 24

US-10-204-708-35
; Sequence 35, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation.
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; CURRENT FILING DATE: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 35
; LENGTH: 9347
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-35

Query Match 2.6%; Score 92.3004; DB 1; Length 9347;
Best Local Similarity 48.3%; Pred. No. 16;
Matches 366; Conservative 0; Mismatches 377; Indels 15; Gaps 4;
QY 2813 TTTATATTACTCAGAAAATAGTAATTTTCATATGATTCACAAATTTATTTCAATAATGTTGGT 2872
Db 8555 TTTTATAGTTTGTGTTGTTATCGTTTTTATAGGTTTTCGATTTATTTTAGAGGTTT 8614
QY 2873 TAAGATAATAGATTTTCAAAATGATTTTATCTTGAATTTTCTCTACTTTATTAATTT 2932
Db 8615 AAGCGTTTAAAGTTAAGTTGGATTTTATTTTATTTTATAGATAGTTTATTTT 8674
QY 2933 TGGGATTTAACTATTCTTCAATGACTGTTATTTCTAATATTACTTATTTCTATTATAC 2992
Db 8675 GCGGATTTATTTGGGGTTTTATAGTTTATATATTTTATGTTTATTTGATGGTTTTTAA 8734
QY 2993 TTTAATTCACHTA-----TTTTTATGATTTTCTAATAAATCCAGTCTCTGTTTTT 3046
Db 8735 TTTATATTTTCTTTTAAATTTTGGTTTGAATATAGTTGGTTGTTTGTAGTTTGTCTT 8794
QY 3047 TTAATAAGACTTTAAATTTATTAATTTCTCTTTTAGTGTTTTACCAGTT--CTTTGAGCT 3104
Db 8795 TTTAGTTGGATGTTTAAATGGTATTTTAAATTTTGTATGTTTAAATTAAGTTTGTGTA 8854
QY 3105 ACTCTTTTGANTTATTGGTCTCTATCTTTTCTCAAGTTTTCAGATTTTGAATGGTAGTAACCA 3164

Db 8955 TTTTTTTTTTTTTTTAGAGTAGTTTTTTTTTGGTTTTTTTTTTTTTTTAGTTAATAGGA 8914
QY 3165 TTTATCTTTATTTTTTTGTAATTAGCTCTTTAAATTCATTATTTCTTTGATTAACAGCTTCA 3224
Db 8915 CGTTATTTTTTTTGTGCTATAGGTTAAAAATATTAGTATTATTTTGTATTTTTTTTTTT 8974
QY 3225 GTTCTATGCGCTTAAATAAGTTTTTTTTTTTTTTTTTTTTTTTTTAAAGAATGCTTTGT 3284
Db 8975 TATATTTATATTTAGTTAGTATTTATTTGTTTATTTTATTTTATGATGATAAGTTTTAT 9034
QY 3285 GAAGTTTTGACAATGCTTTGAGCAATAATTTAGGATATTTTGAATGGTTCATGAGTATG 3344
Db 9035 TTAGAGTTTA-GTTATTTTTCGTTATTTTATTTTATTTTGAATTTAGTTATTATTT 9093
QY 3345 CTTTTGTACTTGGCAATTTATTTGAAGTTTAT-----GATTTATGAATTTATGATGCTTTT 3398
Db 9094 ATTTTTTATTTGGGTTTTTTTATGATTTTTTAAATTTGTTTGTGTTTGTATTTTATTT 9153
QY 3399 TTTGGGCATAAAGGTTCTATGGCATATTTTTTTGTGTTCTATATTTCTTAAATTTATAATTG 3458
Db 9154 TTTTTTGTTTTTTTTGTGTTGAGTTTGTGAAGGATAAGGATGTTTTTGTTTTTTATTT 9213
QY 3459 GCTTTAAAAGTATTTGCTGCTATTAACATCAATTAAGTCTTATTTTGGACTATAGTGA 3518
Db 9214 GTTTAAATAGTGTGTTGTTATTAATAGATTTTGAATAAATATTGATGGAGTAATAAGGA 9273
QY 3519 GTCAACAAAAGAGTTGGACATGACTTTAGGACTTAAGCAA 3556
Db 9274 TTATTTTGTTTTTTTAGGTGAAATTTTCAGAGAAATGAA 9311

RESULT 25

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)833-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZPT-F15
US-08-232-463-14

Query Match 2.6%; Score 91.9999; DB 1; Length 7218;
Best Local Similarity 9.5%; Pred. No. 21;
Matches 44; Conservative 249; Mismatches 169; Indels 0; Gaps 0;
QY 2840 CATATGTAATCAAAATTAATTCATATGTCGTTAAGATAATAAGATTTCAAAATGAT 2899
Db |||||
QY 2900 TTATCTTTGATTTTCTCTACTATTAATTAATTTGGGATTTTAACATATTTCTCAATGAC 2959
Db |||||
QY 1072 YY 1071
Db |||||
QY 2960 TGTATTTCTAATATTTACTTATTTACTTTAATTCACACTATTTTATGATTT 3019
Db |||||
QY 1132 YY 1191
Db |||||
QY 3020 TTCTAATAAATCCAGCTCTTCTTTTAAAGAGACTTTAAATTAATTTCTCTT 3079
Db |||||
QY 1192 YY 1251
Db |||||
QY 3080 AGTGTTTACAGTCTTTCAGGCTACTCTTTGATTTATTTGGTCTCTATCTTCTCA 3139
Db |||||
QY 1252 YY 1311
Db |||||
QY 3140 AGTTTGAATGGCTACGTAATCATTTATCTTTATTTTGTAAATGCTCTTAAAT 3199
Db |||||
QY 1312 YY 1371
Db |||||
QY 3200 TCATATTTCTTGAATACAGCTTCAGTCTATGCTTTAATAAGTTTTTTTTTTTT 3259
Db |||||
QY 1372 YY 1431
Db |||||
QY 3260 TTTTAAAGATGCTATCTTCTGAGTTTTCACATGCT 3301
Db |||||
QY 1432 YYYGTACAAATCTCTATCTCTTAACTACTTGCTAGAT 1473
Db |||||

RESULT 26
US-10-204-708-3
Sequence 3, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 3
LENGTH: 10619
TYPE: DNA
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-3

Query Match 2.6%; Score 91.5004; DB 1; Length 10619;
Best Local Similarity 47.4%; Pred. No. 15;
Matches 332; Conservative 0; Mismatches 365; Indels 3; Gaps 2;
QY 2865 ATGTGCTTAAGATAATAAGATTTTCAAAATGATTTTATCTTCTTCTCTACTTA 2924
Db |||||
QY 3128 ATAGTTGTTAAATTAAGATGATGATGATGATGATGATGATGATGATGATGAT 3187
Db |||||
QY 2925 TTTAAATTTGGGATTTAACTATTTCTTCAATGACTGTGATTTCTAATAATTTACTTATTC 2984
Db |||||
QY 3188 AATTTTATTTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 3247
Db |||||
QY 2985 TATTTTACTTTTAACTGCTATTTTATTTATTTATTTATTTATTTATTTATTTATTT 3044
Db |||||
QY 3248 TTTTATTTGAGATAGGTTATTTATTTATTTAGGTAGAGTTTATTTATTTATTTATTT 3307
Db |||||
QY 3045 TTTTAAAGAGACTTTAAATTTATTTATTTCTTTAGTCTTTTACAGTCTTTTACAGCT 3104
Db |||||
QY 3308 TTTATTTATAGTTTAAATTTTATTTAGGTTTATTTATTTATTTATTTATTTATTTATTT 3367
Db |||||
QY 3105 ACTTCTTTTGTATTTTGGTCTCTATCTTTCTCAAGTTTGAATTTGGTCTAGTAAC 3164
Db |||||
QY 3368 ATGGGATTAATAGTATTTGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 3427
Db |||||
QY 3165 TTTATCTTTA--TTTTTTGTAATTTAGCTCTTTAAATTTCAATTTCTTTGATAACAGCTT 3222
Db |||||
QY 3428 ATGTTGTTTATGCTGTTTGAATTTTGGGTTTAAAGGATTCGTTATTTATTTATTTATTT 3487
Db |||||
QY 3223 CAGTCTTATGCTTTTAAAGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 3282
Db |||||
QY 3488 AAGGTTTATGATTTATAGGTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 3547
Db |||||
QY 3283 GTCAAGTTTGTACATGCTTTTGAAGCAATTTTGAAGTATTTTGAAGTATTTTGAAGT 3342
Db |||||
QY 3548 GGTATTAATTTTATGATTTTATGATTTTATGATTTTATGATTTTATGATTTTATGAT 3607
Db |||||
QY 3343 TGTCTTTGATTTGCTATTTTATGAGTTTATGATTTTATGATTTTATGATTTTATGAT 3402
Db |||||
QY 3608 TGTCTTTTATGATTTTATGATTTTATGATTTTATGATTTTATGATTTTATGATTT 3667
Db |||||
QY 3403 GGCATAAAGGCTTATGCTATTTTATGATTTTATGATTTTATGATTTTATGATTT 3461
Db |||||
QY 3668 TGTCTTTTATGATTTTATGATTTTATGATTTTATGATTTTATGATTTTATGATTT 3727
Db |||||
QY 3462 TTAAGAGTATTTGCTGCTATTTAAACATGATTTAAAGTCTTTATTTGGACTATAGTGG 3521
Db |||||
QY 3728 TTAATTTTGAATGAAATGAAATGATGATGATGATGATGATGATGATGATGATGAT 3787
Db |||||
QY 3522 ACATAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3581
Db |||||
QY 3788 AATTATTTTGAAGTATGATTTTATGATTTTATGATTTTATGATTTTATGATTT 3827
Db |||||

RESULT 27
US-10-204-708-23
Sequence 23, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8

; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 23
; LENGTH: 11049
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-23

Query Match 2.5%; Score 90.4004; DB 1; Length 11049;
Best Local Similarity 47.8%; Pred. No. 15;
Matches 343; Conservative 0; Mismatches 356; Indels 18; Gaps 3;

QY 2806 TTAGTAGTTATATTAATCTACTCAGAAATAGTAATTCATATGATTTCAAAATTTATTCATAA 2865
Db TTTGAAATTTTTTTTAAAGTAAATGCTGATTTATGTTATTTAAATTTATTTAGTTT 8243
QY 2866 TGTGTTAAGATAATAAGATTTTCAAAATGATTTTATCTTGTGATTTTCTCTACTAT 2925
Db TTTT-----TTATTTATGACGTTTTTTTTTTTTTTTTTTTATTTATTTAGAAAGTTT 8293
QY 2926 TTAATTTGGGATTTAACTATTTCTCAATGACTTGTATTTCTAAATTTTACTTATCT 2985
Db TTTAAATTTTTTATTTTAAATTTGTTTTTATATTAATAATATTGTTTTTATGATTT 8353
QY 2986 ATTTACTTTAATTCACATTTATTTATGATTTTCTAAATCAATCCAGTCCTTGT 3045
Db ATTTATAGAAATAGAAATAGGTTATTTAATTTTGGTGTGAATGGTATTTTTTTTT 8413
QY 3046 TTTAAAAGACTTTAAAATTTAATTTCTCTTTAGTGTTTTACCAGTCTTTCCAGCTA 3105
Db TTTAAAAGAAATTTTATGCTGTATAAATAAGAAATGAAATAGATTTGTTAATG 8473
QY 3106 CTTCTTTGATTTATTTGGTCTCTATCTTTCTCAAGTTTTTGAATGGCTAGTAACTCAT 3165
Db TTAGTTAGTATAGATTTTGGTGGAAATTTGTTTTTTTGGTGGAAATGAGGTGAG 8333
QY 3166 TTATCTTTATTTTGTAAAT-----TAGCTCTTTAAATTCATTATTTCTTTGATAACAG 3219
Db TTGAGTTTTTTTTTGAATAATATTTAGTAGGTTTTTAAAGAGGAAATTTTTTTATTTGTT 8593
QY 3220 CTTCACTCTATGGCTTTAATAAGCTTTTTTTTTTTTTTTTTTTTTTAAAGAACTCA--T 3277
Db TGTATTTTTTAAATTTTTTAAAGTATAGTTTTTTTCGTAATTTTTTAAAGATGATTTAGGT 8653
QY 3278 TCTTTGTGAAGTTTTGACAATGCTTTGAGCAATATTTAGGATATTTTGAATGGTTCAAT 3337
Db ATTTGTTAGAGTTTTTGGGAATTTTGGAAATTTGTTATTTTGTGATGTTTATATGTTAT 8713
QY 3338 GAGTATGCTTTTGTACTTGGCAATTTATGAAGTTATGATTTATGAATATGATGCTTTT 3397
Db TTAGAATTTTTTTTGGAGTTAGATTTGATTTTATTTTGAATTTGATTTGTAATGTTATGAT 8773
QY 3398 TTTTGGCCATAAAGGCTATGSCATATTTTGTGCTATATCTTAAATTTATAAAT 3457
Db TATATGTAAGTTGTTGTAGTTTTGTTTTTTTATTTGTTTTTTTTTTCAGAGTTTTATTT 8833
QY 3458 GGCCTTAAAGATTTTGGTGTATTAACATGAATTAAGTCTTATTTTGGCTATAG 3514
Db TTGATGTAATTTATTTGGTGAATTTATTTGTTTGAATTTAAATTTTTTTTATTTATTTAG 8890

RESULT 28
US-10-204-708-62
; Sequence 62, Application US/10204708
; Patent No. 6677731

GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: FIEBENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: By Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 62
; LENGTH: 6801
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-62

Query Match 2.5%; Score 89.7004; DB 1; Length 6801;
Best Local Similarity 47.1%; Pred. No. 24;
Matches 304; Conservative 0; Mismatches 338; Indels 3; Gaps 1;

QY 2863 TAATCTGGTTAGATATAAGATTTCAAAATGATTTTATCTTCTGATTTTCTCTACT 2922
Db 5606 TTAATTTTGTAAATAATTTTTTTTTTTAGTTTTTTTTTATTTATTTATTTT 5665
QY 2923 TATTTAAATTTTGGGATTTTAACTATTTCTCAATGACTGTGATTTCTAAATTTACTTAT 2982
Db 5666 TTTTTTCGGTTTCGTTTTTTTTTATTTTTTTTGGTTTTTTTATTTTAAATTTTTTTTTTT 5725
QY 2983 TCTATTTTACITTAATTCGACTTATTTTATTTATTTGATTTTCTAATAAATCCAGTCCTTGT 3042
Db 5726 TATGTTATATTCGGTTTTGTTTAAATTTTTTTTTTTGTTTAAAGGTTATTTTATTTATTTAAA 5785
QY 3043 TTTTAAAAAGACTTTAAAAATTTAATTTCTCTTTAGTGTTTTACCAGTCTTTCCAGG 3102
Db 5786 GGTTTTATAAAGTTTTTTTTTTGAGAAAGT---GTTAGGTCGGAAGGTATAGTTTAAGT 5842
QY 3103 CTACTCTTTTGAATTTATTTGTCCTATCTTTCTCAAGTTTTTGAATTTGGCTAGTAACT 3162
Db 5843 TTATGTTTTTTTTTTTATTTAGTTAAATTTTTTTTATTTTAAAGAGATTTAGTTTTTTTTT 5902
QY 3163 CATTATCTTTATTTTTTGTAAATTAGCTCTTTAAATTCATTATTTCTTGTATACAGCTT 3222
Db 5903 TATCGATTTTAGTTTTTTTTTATAAGGAATTTTGTATGTTATTTTATGATTTTAT 5962
QY 3223 CAGTTCTATGGCTTTTAAAGTTTTTTTTTTTTTTTTTTTTTTTTTAAAGATGTCATCTTT 3282
Db 5963 TTGGTATGACGTTATGTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 6022
QY 3283 GTGAAGTTTTCACATGCTTTGAGCAATAATTTAGGATATTTTGAATGGTTCTATGAGTA 3342
Db 6023 TTAGATAATTTATTTAGTTTATGATATTTTATTTTATTTATTAAGATCTGTTTATATTT 6082
QY 3343 TCGTTTGTACTTGGCATTTTATGAAAGTTTAAAGTTTATGATTTTATGATTTATGATGCTTTT 3402
Db 6083 TATAGGGATTTTATGTTGTTTTTAAAGATTTGATTTTAAATTTGGAATTTATGAAAGTAGTTT 6142
QY 3403 GGCATAAAGGCTATGCGCATTTTTTGTGCTATATTTCTTAAATTTATAAATTTAAATTTGGCT 3462
Db 6143 AGATAGATAGAAACGAAGTTATATGATTTATTTTATTTTATTTTATTTTATTTTATTTT 6202
QY 3463 TAAAAAGTATTTGCTGCTATTTAAACATGAATTAAGTCTTTATTTG 3507

Db 784 TTTTGTGAAGCGTTTATTTA--TATATATATATATATATATATATATATATAT 841
QY 3226 TTCTATGGCTTTAATAAGATTTTTTTTTTTTTTTTTTAAAGAAATGTCATCTTTGTG 3285
Db 842 ATATATTTTTTTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 901
QY 3286 AGTTTGTGACAACTTTGAGCAATATTTTGAATATTTTGAATGGTTTCATGATGTC 3345
Db 902 AGTTTGTGTTTTTTTTTAAATTTTGTGGGTTTTTATTTGTTTTTTTATTTAGTTCGT 961
QY 3346 TTTTGTACTTGGCAATTTAT-----GAAGTTTATGATTTATGAATTAATGATGCTT 3395
Db 962 TTTTAAATTTTATTTAGTTTATTTTATTTTGGGAATTTTTTTTTTTTTTTTTTTTTT 1021
QY 3396 TTTTGTGGCAAAAGCTATGCGCATATTTTGTGGTCTATATTCITTAATAATATATA 3455
Db 1022 TTAGTTTGTAGAAATTTGATTTTATATAATTTTGTGGGTTTATTTTATTTAGAAATTTAT 1081
QY 3456 TTGGCTTTAAAGATTTTGTCTCTATTAAACATGAATTAAGCTTTATT 3504
Db 1082 TATTTGTTAAAGTTAAAGTTTTTTTTTTTTTTTTTTTTTAAATTTTAAATTTTAAAT 1130

RESULT 31

US-08-998-416-1137
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtie, Philipp
; APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 623926artis Corporation
; STREET: 3054 Cornwalis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CSC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:

ORGANISM: PAG1692RP
US-08-998-416-1137
Query Match 2.4%; Score 85.4004; DB 1; Length 636;
Best Local Similarity 48.2%; Pred. No. 2.7e+02;
Matches 295; Conservative 0; Mismatches 311; Indels 6; Gaps 2;
QY 2843 ATGTATTCAAAATTTATTTCAATGTTGGTTAAGATAAATGAATTTTCAATGATTTT 2902
Db 22 ATTATATAAGATTAAATAAATTTTATTTATATAATTTTAAAGATTAAATTTAAAC 81
QY 2903 AACTTGTATTTTCTCTACTTATTTAAATTTTGGGATTTTAACTATTCTTCAATGACTTG 2962
Db 82 TATTATTATCATTTATTAATAAATTTATTTTCAATTTATTAATTAATTAATTAATTA 141
QY 2963 TATTCTTAATATTTACTTATTTCTATTTTACTTTAAATTTGCACTTATTTTATTTGATTTTC 3022
Db 142 TATATAATTTACTTAATTCATCATTAATTAATTTTATATAATTAATAAATAATTAATA 201
QY 3023 T---AATAAAATCCAGTCCCTGTTT---TTTTAAAGAGACTTTTAAATTTATTAATTTCTC 3076
Db 202 TATGAATACTATTAGTCTATGTTCAAAATTTTAAATTTAGTTTATTAATAATTTATTAGATA 261
QY 3077 TTTAGTGTTTTACCAGTTCTTTTCCAGGCTACTCTTTTGATTTTATTTGGTCTCTATCTTTTC 3136
Db 262 TTATTATTTTCTTTAAATAAATTTATTAATAGATTATCAATTAATTAATTAATTAATTA 321
QY 3137 TCAAGTTTGTGAATTTGGCTACGTAATCTTATTTATTTTGTGAAATTTAGTCTTTTA 3196
Db 322 ATTGTTTATTAATAATAATTTATTTATTAATAAAGATTTTAAATTTATTTAAATATTGTA 381
QY 3197 AATTCATTTATTTTGAATAACAGCTTCAGTTTCATGCTGTTTAAATAAGCTTTTTTTTTTT 3256
Db 382 AATTATTTATTTTATTAATAATCTTATTTTATAAATAATTTATTTGATTTTATTTATTA 441
QY 3257 TTTTTTTTTTAAAGAATGTCATTTCTTTGTGAAGTTTGTGCAATGCTTTGAGCAATAATTTA 3316
Db 442 ATCTTTTATAGAATTTATTTATAAATTAATTTTAACTTTTAAATTTCTTATTTAAATTT 501
QY 3317 GGATATTTTGTGAATGGTTCATGAGATATGCTTTGTGACTTTGGCAATTTATGAATTTATGA 3376
Db 502 TTATATTATTTAAATAAATTTATTTCAATTTTATTTTATTTTATTTTAAATTTAAATTA 561
QY 3377 TTTTATGAATTTATGATGCTTTTGTGGGCAATAAAGGCTATGCGCATATTTTGTGGTCT 3436
Db 562 TTTAAATTAATTTTATCATTAATTTTAAATAAATAATTAATAAAGATGTAGTTA 621
QY 3437 ATATTTCTTAAAA 3448
Db 622 AAAATACTTATA 633
RESULT 32
US-10-204-708-14
; Sequence 14, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 14
 ; LENGTH: 6113
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-14

Query Match 2.4%; Score 84.5003; DB 1; Length 6113;
 Best Local Similarity 48.9%; Pred. No. 31;
 Matches 303; Conservative 0; Mismatches 300; Indels 17; Gaps 3;

QY	2811	AGTTTATTAATCACTCAGAAAATAGTAATTCATATGATTCACAAATTTATTCATAATTTTTCATAATCTG	2870
DB	46	AATTATTAAGTAATGCTGTTTTTAAGTAATATTTTATAGGTTTTTAAATTAATGCG	105
QY	2871	GTTAAGATAATAAGATTTTCAATGATTTTATCTTTGATTTTCTCTACTTATTTAAT	2930
DB	106	GTAATAATTAAGATATAATTTGTTTTTTTTTTCGAAAATTTTTTTTA--GAGTATAT	163
QY	2931	TTTGGGATTTTAACTATTCTTCAATGACATGTAATTTCTAATATTTACTTATTTCTATTTT	2990
DB	164	TATTAATATTTAATTTTATATAGGATAGATTGCGGAATTAATATTTATTTATTTTTTT	223
QY	2991	ACTTTAATGCACTAATTTTATGATTTTCTAATAAATCCAGTCCTTTGTTTTTTTAA	3050
DB	224	TTTTGTTATTTACGTAATTT-----ATTAAATTTTATGTTGTTGATTTTATTTTAA	277
QY	3051	AAAGACTTTAAATATATATTTCTTTAGTGTTTTACAGGTCTTTTCAGGCTACTTCT	3110
DB	278	GACGTGTTTAAATTTTTTTTATTAATTTTATTTTATTTGTAATTTTGTGTTTAGGTTT	337
QY	3111	TTTGATTTATTTGGTCCTATCTTTCTCAAGTTTGTAAATGGGTACGTAATCTCATTTATC	3170
DB	338	TTAATAATTTTTCGGAT-----AAGTGAATTAATAATAGTATTTTAATTTGGT	388
QY	3171	TTTATTTTGTAAATAGTCCTTTAAATCAATTTCTTTGTATGACAGCTTCAGTTCTA	3230
DB	389	TTTTATTTTGTGTTTATGTTAAATTTATTTTAAAGTAAATAAAATGTAATAGTTG	448
QY	3231	TGGCTTTAATAAGTTTTTTTTTTTTTTTTTTTTTTTAAAGATGTCATTTCTTTGTGAAGTT	3290
DB	449	TTATTTTATTTTATTTTATATTTTATTTTAAAGTATATAAGGTTTTTTTTTATTTAGTT	508
QY	3291	TTGACAATGCTTTGAGCAATAATTTAGGATATTTTGAATGGTTTCATGATGCTTTTGG	3350
DB	509	TTTGTTTATTTTATTTTGTGTAATGTTTTTATTTAAATTTTTCGTTAGTATTTTGG	568
QY	3351	TACTTGCCATTTATGAAATTTATGATTTATGAATATGATGCTTTTTTTTGGGCATAAA	3410
DB	569	AATTTGTTTAAATTTTGGTTGTTGTTTTTTGTTTTTAAATAGGTTGTTTGGTATATTGA	628
QY	3411	GGTCTATGCAATATTTTTTGG 3430	
DB	629	GGAATTTAATAATATGATG 648	

RESULT 33

US-10-204-708-31
 ; Sequence 31, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; CURRENT FILING DATE: 2003-05-06

; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 31
 ; LENGTH: 8093
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-31

Query Match 2.3%; Score 83.8002; DB 1; Length 8093;
 Best Local Similarity 48.3%; Pred. No. 24;
 Matches 262; Conservative 0; Mismatches 277; Indels 4; Gaps 1;

QY	2848	TTCAAAATTAATTCATTAATGTTGGTTAAGATAATAAGATTTTCAAAATGATTTTAT---	2904
DB	3123	TTTTAAAATGTTTTTTTAAATTTTTTTTGGAGACGATAGATTATTATTGATATTATTAG	3182
QY	2905	-CTTTGATTTTCTCTACTTATTTAAATTTTGGGATTTTAACTATTTCTTCAATGACTGT	2963
DB	3183	GATTAATATTTTATTTTATTTTGTGTTTATTTAGTTTATTAATAATTTTGTGA	3242
QY	2964	ATTTCTAATATTTACTTATTTACTTTAATTTGCACTTATTTTATTTGATTTTCT	3023
DB	3243	GGGTAGGAAAGTTTGTAGATTTTTTTTTTTAAATGTTTGTGAAATAAATTTAGTTTAA	3302
QY	3024	AATAAATCCAGTCCTGTTGTTTTTTTAAAGACCTTTAAATATTAATTTCTTTTAGTG	3083
DB	3303	TTTCGAAAGTATAAATATGGTTAAAGGTAGTAGTAAGATTTTATTAATTCGTAAT	3362
QY	3084	TTTTACCAAGTTCTTTTCAGGCTACTTTTTCATTTATTTGGTCCTATCTTTCTCAAGTT	3143
DB	3363	TTTTATTGAGTTTTCGAAAGGTGTTTTATAATGTTTATTAAGAGCTTAAAGTTATAAGTT	3422
QY	3144	TTGAATGGCTACGTAATCTATTTATTTATTTTGTGTAATTTAGCTCTTTAAATTCAT	3203
DB	3423	ATTGTTGTTTGAAGAAGTGGATTGTTGTTTATTTGAAATTAAGACGTTTATTT	3482
QY	3204	TATTTCTTTGATAACAGCTTTCAGTTCTATGCGTTTAAATAAGTTTTTTTTTTTTTTT	3263
DB	3483	TTTTTTTAGAGAGAATAATAGTTTAAAGTTTTTTTTTTTAAAGTTAATTTAGAT	3542
QY	3264	TAAAGAATGCAATCTTTGTGAAGTTTTCACAAGCTTTGAGCAATTAATTTAGGATATT	3323
DB	3543	ATCGTTTATTAATTTTTTTTATTTATGTTTGTAGATTTCGTTAGTAAAGTTTGAAT	3602
QY	3324	TTTGAATGGTTTCATGAGTATGCTTTTCTACTTGCATTTTATGAGTTTATGATTTATGA	3383
DB	3603	TAGTCTGTTTAACGTTTTTGTGTTTAAATAATTTATTTATTCGGATATGGTATTATGA	3662
QY	3384	ATT 3386	
DB	3663	TTT 3665	

RESULT 34

US-07-867-106-2
 ; Sequence 2, Application US/07867106
 ; Patent No. 5389526
 ; GENERAL INFORMATION:

; APPLICANT: Slade, Martin B
 ; APPLICANT: Chang, Andy C M
 ; APPLICANT: Williams, Keith L
 ; TITLE OF INVENTION: Improved Plasmid Vectors for Cellular

TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5399526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2

Query Match 2.3%; Score 82.6003; DB 1; Length 5852;
Best Local Similarity 46.1%; Pred. No. 33; Mismatches 369; Indels 22; Gaps 2;
Matches 335; Conservative 0;

QY 2796 TACTGAAACCTTAGTGTATATATCTCAGAAATAGTAATTTTCATATGATATTCATAAT 2855
DB 1772 TATCGATATACITTAATTTATTAAGATGGAATATATTTTAAATTAATAATTCATC 1831
QY 2856 TATTTCATAATGTTGGTTAAGATAATAAGATTTTCAAAATGCAITTTTATCTTTGATTTT 2915
DB 1832 TTTTTTTTTTTTCTTTTTTTTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1891
QY 2916 CTCTACTTATTTAATTTTGGGATTTTAACTATTTCT-----TCAATGAC 2959
DB 1892 TAATTAATAATCAATTTAAATTAAGTAATTTTAAATTAAGTAATTTTAAATTAAGTAATTT 1951
QY 2960 TTGTAATTTCTAATATTTACTTATTTTACTTATTTTAAATTTTAAATTTTAAATTTTAAATTT 3019
DB 1952 TAAAGTAGTTTAAACCAACTTTTCTATTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTACTTT 2011
QY 3020 TTCTTAATAAATCCAGTCTGTTTCTTTTAAAGACCTTTTAAATTTTAAATTTTAAATTTTCTCTTT 3079
DB 2012 GAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2071
QY 3080 AGTGTTTACCAGTCTCTTTCAGGCTACTCTTTTGAATTTTATTTGGTCTATCTTTTCTCA 3139
DB 2072 TTTTGTGATTTTCTTTTAAATAAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 2131
QY 3140 AGTTTGAATGGCTACGTAATCTATTTATCTTTATTTTGTAAATAGCTCTTTAAAT 3199

DB 2132 TTTATAAATATTGGATAATATATCAAAATATTTATCAGTTTGGCATGCAATTTTATTA 2191
QY 3200 TCATTATTTCTTTGATAACAGCTTCAGTTCTATGGCTTTAATAAAGTTTTTTTTTTTTTT 3259
DB 2192 TATTTATTTTGTATTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2251
QY 3260 TTTTAAAGAAATGTCATTTCTTTGTGAAGTTTTCACATGCTTTTGACCAATTAATTAGGA 3319
DB 2252 TTTTAAATTTTAAATTTTAAATTTTCCACACTTTTCATTTTATTTTATTTTATTTAT 2311
QY 3320 TATTTTGAATGGTTTCATGAGTATGCTTTTCTACTTGGCATTATTTGAAGTTTATGATTT 3379
DB 2312 TGTAAATTCATTTATTTTATTTTAAATAGTTTGTGTTTAAATTTTATTTTCAAGATT 2371
QY 3380 ATGAATATGATGCTTTTGTGGGCATAAAGGCTCTATGGCATATTTTGTGGTCTATA 3439
DB 2372 TTAATAATGGACGAACTTTATTTCTTGGGATAGGTTTTTAAAGTTTTTTTGTATATA 2425
QY 3440 TCTTAAATTAATAATTTGGCTTTAAAGAGTATTTGCTGCTATTAAACATGAATTAAGTC 3499
DB 2426 CTTTGGAGAAATTCAAAGTTGTAAAGAAATGATGCGTTTGAGTGTGATTTATGAC 2485
QY 3500 TTATTT 3505
DB 2486 ATTCTT 2491

RESULT 35
US-10-204-708-86
; Sequence 86, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: FIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: By Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 86
; LENGTH: 11050
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-86

Query Match 2.3%; Score 82.4003; DB 1; Length 11050;
Best Local Similarity 46.5%; Pred. No. 18; Mismatches 306; Indels 0; Gaps 0;
Matches 266; Conservative 0;

QY 2923 TATTTAATTTGGGATTTTAACTATTTCTTCAATGACTTGTATTTCTAATATTTACTTAT 2982
DB 7114 TTTTATGTTTGGGTTTAAAGTCGATTTTAAAGTATTTTATTTAGATTGAGTTT 7173
QY 2983 TCTATTTTACTTTAATTTGCACTTATTTTATTTGATTTTCTAATAAATCCAGTCTTCT 3042
DB 7174 TATATTTAAATTTATTTATGAGTTTATTTATTTAGTTTAAAGATGAGTTAATTTT 7233
QY 3043 TTTTAAAAAGACTTTTAAATTTTAAATTTTCTTTTAGTGTTTTACCAGTCTTTTTCAGG 3102

```

Db 7234 TTAATTTTATTTAAGGATTTTCGAGGAAATTAATCGGTATTATAGATTGCGTTTCGGA 7293
QY 3103 CTACTCTCTTTGATTTATTTGGTCTCTATCTTTCTCAAGTTTTCGAAATGGCTAGCTAACT 3162
Db 7294 TTTAAGGATTTTATTTTGTAAATCGTTAGGTGTTTATGATTTGAGTTAGTTTGGTAAAT 7353
QY 3163 CATTTATCTTTATTTTGTGAAATAGCTCTTTAAATTCATATCTCTTGATTAACAGCTT 3222
Db 7354 TTTGTTTTCGGGTTTAAATATATAGTTTGGGTTTTTTTATATATATTTTGTAGTTTAA 7413
QY 3223 CAGTTCTATGCTTTAAATAAGTTTTTTTTTTTTTTTTTTTTTTTTTAAAGATGTCATCTTT 3282
Db 7414 GATTTTGAATTTAGTTTAAATTTTAAATTTTAAATTTTATTTTGGAGATCGTTTTCGTTG 7473
QY 3283 GTGAAGTTTTCGAAATGCTTTTGACAAATAATTTAGGATATTTTGAATGTTTCATGAGTA 3342
Db 7474 GAGGATTTTATTTTTCGTTTTTGAATTTTAAATATTTTGTAGATTTTAAATGA 7533
QY 3343 TGCCTTTTGTACTTCGCATTTATTTGAAGTTTATGATTTTATGATGCTTTTTCGTTG 3402
Db 7534 TTTTATTAAGTTTTTTTTTTTAAATTAATTAATTTTAAATTTTGTAGTTT 7593
QY 3403 GGCATAAGGTTCTAGGCATATTTTTTGTGCTCTATATCTTAAATTAATAAATGCGCTT 3462
Db 7594 AAGTTAGTAGTATTTATTTTATTTTATTTTAAAGTTTTCGATTTTGTAGTTTAT 7653
QY 3463 TAAAAAGTATTTGCTGCTATTAACATGAAT 3494
Db 7654 TATTTTTTTTGGAGTTAGTAGATTTTGTGT 7685

RESULT 36
US-10-204-708-55
; Sequence 55, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 55
; LENGTH: 11015
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-55

Query Match 2.3%; Score 81.3004; DB 1; Length 11015;
Best Local Similarity 46.7%; Pred. No. 19;
Matches 316; Conservative 0; Mismatches 347; Indels 13; Gaps 2;

QY 2813 TTTATATTACTAGAAAATAGTAAATTCATATGATTAATCAAAATTTATTCATATGTTGT 2872
Db 9215 TATAGTTAGGGAATTTGTTTATTAATGCTGTTATTTATTTATTTAGTTTGTAGA 9274
QY 2873 TAAGATAATAAGATTTTCAAAATGATTTTATCTTT-----GATTTTCTCTACTATT 2927

```

```

Db 9275 AGGTTTTTTTTTTTATTTAAGAGGAAGTTTTTATATTTAAGTAGTATTTTTTTTATTT 9334
QY 2928 AATTTTGGGATTTTAACTATTTCTTCAATGACTGCTGATTTCTAAATATTACTTATCTAT 2987
Db 9335 TTTTATTTAGTTTTTTTAAATTAAGAAATTTGTTTTTTTTTTTATCGAATTTGTTGTTTG 9394
QY 2988 TTTACTTTAAATTCGCACTTATTTTATTTGATTTTCTCAATATAAAATCCAGTCCCTGTTTT 3047
Db 9395 GGTATTTTATAAGTCGATTTATATTTATATGTTTTTTTGTGTAGGTTTTTTTTTAT 9454
QY 3048 TAAAGAAGCTTTAAATTAATTAATTTCTTTAGTTGTTTACCACTTCTTTCAGGCTACT 3107
Db 9455 TAATATAATTTTGAAGTTTATTTATGTTGATTAATGATTAATATATAGTTTTTTTATTTT 9514
QY 3108 TCTTTTGATTTATTTGGTCTATCTTTTCTCAAGTT-----TTGAATTTGGCTACGTA 3159
Db 9515 TTTTTCGTTTTTTCGAGAAGGAGTTTTTTTTTATTCGTTTGGTTGGAGTGTAGTGATG 9574
QY 3160 ACTCATTTATCTTTATTTTTCGAAATTTAGCTCTTTAAATTCATTTCTTTGATAACAG 3219
Db 9575 AGATTCGGTTTTATTTGTAATTCGTTTTTTTAGGTTTAAAGCAATTTTTCGTTTGTAGTT 9634
QY 3220 CTTTCAGTTCTATGGCTTTAAATAAAGTTTTTTTTTTTTTTTTTTTTTTTAAAGATGTCATC 3279
Db 9635 TTTAAGTAGTTGGGATTTAGGATTTATTTATTTAGCTTTAGTTAAATTTTGTATTTTA 9694
QY 3280 TTTGTGAAGTTTGTGCAATGCTTTGAGCAATAATTTAGGATATTTTGAATTTGTTTCATGA 3339
Db 9695 GTAGAGATAGGGTTTTTATTTATGTTGGTTAGGTTTGTGTTTAAATTTTCATTTTCGTTATC 9754
QY 3340 GTATGCTTTTGTACTGTCATTTATGAAAGTTTATGATTTATGATTTATGATGCTTTTTT 3399
Db 9755 GTTTATTTTATGTTTTTAAAGTTTGGAAATTTAGGTAAGATTTAGTTATTTGTTTTT 9814
QY 3400 TTGGGCATAAAGGCTCATGGCATATTTTTTGTGCTCTATATTTCTTAAATATATAAATGG 3459
Db 9815 TTATTTTTTTTATGTTGTAATAATATTTTATTTATATGATAGATATATATTTTATTTAT 9874
QY 3460 CTTTAAAGATTTTG 3475
Db 9875 TATATATTAGTTGATG 9890

RESULT 37
US-10-204-708-45
; Sequence 45, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 45
; LENGTH: 19233
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```



```

/ OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (76, 178..179, 273, 586, 648, 651, 920, 1014, 1173, 1197, 1228)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1264, 1267, 1559, 1607, 1677, 1703, 1742, 1848, 1850, 2033)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (2168, 2254, 2257, 2280, 2291, 2723, 3018, 3080, 4205, 4209)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (4216, 4225, 4239, 4246, 4254, 4265, 4270, 4274, 4288, 4294..4295)
/ OTHER INFORMATION: n is a cr g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (4298, 4363, 4610, 4627, 4675, 4693, 4698, 4738, 4777, 4781)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (4787, 4825, 4829, 4838, 4855, 5462, 5494, 5496, 5500, 5674)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (5690, 5697, 5705, 5715, 5890, 5910, 5913, 6027, 6080, 6109)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (6244, 6288, 6776, 6873, 6891, 6980, 7028, 7656, 7876, 8030)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (8059, 8075, 8632, 8638, 9700, 9717, 9815, 9818, 9867, 9879)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (10005, 10075, 10150, 10159, 10165, 10356, 11123, 11166, 11360)
/ OTHER INFORMATION: n is a cr g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (11703, 11933, 12050, 12340, 12981, 13564, 13764..13765, 13787)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (13846, 44168, 16566, 17006, 17217, 17650, 17653, 17656, 17898)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (18130, 18143, 18151, 18188, 18498, 18549, 18936, 19165)
/ OTHER INFORMATION: n is a or g or c or t
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (19200)
/ OTHER INFORMATION: n is a or g or c or t
/
/ US-10-204-708-45

```

	Query Match	2.3%;	Score 81.3003;	DB 1;	Length 19233;
	Best Local Similarity	48.1%;	Pred. No. 11;		
	Matches 325;	Conservative 0;	Mismatches 335;	Indels 17;	Gaps 4;
Qy	2804	CCTTAGTAGTTTATATTACTCAGAAAATAGTAAATTCATATGATGATTCAAAAATATTTCAT	2863		
Db	5799	CGTTATTATATTCGGTATATTTTGTGTTTTATTAGAGATCGGGTTTTATTATGTTGCT	5858		
Qy	2864	AATGTTGGTTAAGATAATAAGATTTTCAAATGATTTTTTAATCTTCGATTTTTCTCTACAT	2923		
Db	5859	TAGTTTGATTTTCGAATATTAGATTTTAAAGTTNATTCGTCGTTTCGCTTTTNTTNAAAAT	5918		
Qy	2924	ATTTAAATTTGGGATTTTAACTATTTCTTCCTCAATGACCTTGATTTTCTAAATATTACTTAT	2983		

Db	5919	GTTCGAATTATAGGTAGTAAT	-----ATCGCGTTTGGTTTTAAGTAATTTAT	5969
Qy	2984	CTATTTACHTTAATATGCACTTA	TTTTTTATGATTTTTCTAATAAAATCAAGTCCTTGTT	3043
Db	5970	GTTATTTTAAATATATGTTAGTT	TGTTTATAAAATTTTTTTGTATAGAGTTATTTTANNT	6029
Qy	3044	TTTTTAAAAAGACTTAAAAATTA	TAATTTCTCTTAGTCTTTTACCAGTCTTTTCAGGC	3103
Db	6030	TTAAAAAGATAAATTTTCGTTAT	TGTTTGTATTTATTAATGTTTTTAGAGGNTTTTAAGTT	6089
Qy	3104	TACTTCCTTTTGATTTA--	TTTGGTCTATCTTTTCTCAAGTTTGAATTTGGCTACGTAA	3160
Db	6090	TAGTATAGGTATTTTAAAGNTAA	ATTTTTTATTTTTTAAAGTTTGGGTTTTTTTGATTTTT	6149
Qy	3161	CTCATTTATCTTATTTTTTGTAA	ATTTAGTCTTTTAAATTCATTTCTTTGTATTAACAGC	3220
Db	6150	ATTTTACGNGTGTATATATTTAT	TATTAGTTAGTTTAAITTAGAAATGAGAAATTA	6209
Qy	3221	TTGAGTTCTTAGGGCTTTAAATA	AGTTTTTTTTTTTTTTTTTTTTTTTTTAAAGAAATGTCA	3277
Db	6210	TTTTTTTTTATCTTTTATTTATTT	TGGGTATTAGNTTTTTTACGAAGTGAAGTGTATTTT	6269
Qy	3278	-TCCTTTGGAAGTTTTCACAAAT	CTTTTGACAAATAATTTAGGATATTTTTTGAAATGGTTCA	3336
Db	6270	GTTTTTGGTAGATTATNTAAGTT	ATTTGGTTTTTTTTTTTATTTATTTATTCGTATCGTTTT	6329
Qy	3337	TGAGTATGCTTTTCTACTTGCA	TTTATTTGAAGTTTATGATTTATGATATGATGCTTT	3396
Db	6330	TTTTTTTTGTTTGAGTCTGTGAT	TTTTTTTATTTAGTATTTTATTTATTTATTTTATTTT	6388
Qy	3397	TTTTTTGGGCATAAAGGCTATGG	CATATTTTTTGTGGTCTATATCTTAAAAATATATAAT	3456
Db	6389	ATAGTATATTTTATTTTATTTT	TTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	6448
Qy	3457	TGGCTTTAAAAAGTATTT	3474	
Db	6449	TTGGTATATTTTATTTT	6466	

RESULT 38
 US-08-213-419B-3/c
 ; Sequence 3, Application US/08213419B
 ; Patent No. 6333406
 ; GENERAL INFORMATION:
 ; APPLICANT: Inselburg, J. et al.
 ; TITLE OF INVENTION: GENE ENCODING PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
 ; TITLE OF INVENTION: AND USES THEREFOR
 ; FILE REFERENCE: J11-002CNCP
 ; CURRENT APPLICATION NUMBER: US/08/213,419B
 ; CURRENT FILING DATE: 1994-03-14
 ; PRIOR APPLICATION NUMBER: US 07/870,506
 ; PRIOR FILING DATE: 1992-04-17
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 6124
 ; TYPE: DNA
 ; ORGANISM: Plasmodium falciparum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (2407)..(2439)
 ; NAME/KEY: CDS
 ; LOCATION: (2598)..(3404)
 ; NAME/KEY: CDS
 ; LOCATION: (3580)..(3720)
 ; NAME/KEY: CDS
 ; LOCATION: (3850)..(5835)
 ; US-08-213-419B-3

```
Query Match      2.3%; Score 81.0002; DB 1; Length 6124;
Best Local Similarity 47.2%; Pred. No. 33;
Matches 246; Conservative 0; Mismatches 275; Indels 0; Gaps 0;
```

QY 2802 AACCTTAGTGTATATATTACTCAGAAATAGTAATTCATATGTAATTCATAAATTAATTC 2861
DB 2621 AACATTTTGTAAATATAACACCTAAABAAATATAGCACAATGTTGTGTATGCGTA 2562
QY 2862 ATATGTTGGTAAAGATTAAGATTTTCAAAATGATTTTATCTTTGATTTTCTCTAC 2921
DB 2561 AAAATATTATATATTTATTTATATAAATATATATTTTACATATTTGTTTCTTTTC 2502
QY 2922 TATTTAAATTTGGGATTTTAACTATTTCTCAATGACTTCTATTTCTAATATTACTTA 2981
DB 2501 TTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 2442
QY 2982 TCTATTTTACTTTAATGCACTTATTTTATTTATGATTTTCTAATAAATCCAGTCTTG 3041
DB 2441 CACAATATGAAAAACAAGGAAATATATGACTTCTATTTTGATATATGAATATTAAGTTATA 2382
QY 3042 TTTTAAAGAGACTTTAAATATTAAATTTCTTTTAGTGTATTTACCAGTCTTTTCAG 3101
DB 2381 TAAATTTTAAAGACTTTTAAATTTTGAATATGTTTGAATTTTAAATTTTCTTTATG 2322
QY 3102 GCTACTCTTTTGATTTATTTGGTCTATCTTTCTCAAGTTTGAATTTGGCTACGTAAC 3161
DB 2321 TTATAATATATGATTTTATAATATATATATATATATTTTATTTTATTTTITTTTGA 2262
QY 3162 TCATTTATCTTTTATTTTGTAAATAGCTCTTTAAATTCATTTCTTTTGATTAACAGCT 3221
DB 2261 CAATCTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 2202
QY 3222 TCAGTCTATGCTTTTAAATAGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 3281
DB 2201 TAAATATATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 2142
QY 3282 TGTGAAGTTTGCATGCTTTGAGCAATTAATTTAGGATAT 3322
DB 2141 TTTTGTGATCTCTCATAAGAGATAAATCTAGAATAT 2101

RESULT 39

US-10-204-708-38
; Sequence 38, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR FILING DATE: 2001-04-06
; PRIOR FILING DATE: 2000-04-06
; PRIOR FILING DATE: 2000-04-07
; PRIOR FILING DATE: 2000-04-07
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; LENGTH: 5501
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-38

Query Match 2.3%; Score 80.4004; DB 1; Length 5501;
Best Local Similarity 47.1%; Pred. No. 38;
Matches 334; Conservative 0; Mismatches 361; Indels 14; Gaps 3;

QY 2806 TTAGTAGTTTATATATCTCAGAAATAGTAATTCATATGTAATTCATAAATTAATTCATAA 2865
DB 1747 TTAATTTATTTTGGAGGATTAAGTTTAAATAAATTTTATGTTTGTAAATTTATTTT 1806
QY 2866 TGTGTTAAGATAATAAGATTTTCAAAATGATTTTATCTTTGATTTTCTCTACTTAT 2925
DB 1807 GATTTAGTTTAAAGTTTGTAGTGTAGTGAATTTTATTTTATTTTATTTTATTTTAT 1866
QY 2926 TTAATTTTGGGATTTTAACTATTTCTCAATGACTTGTATTTCTAATATTACTTATCT 2985
DB 1867 ATATAGATTTGTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTT 1926
QY 2986 ATTTTACTTTTAAATGCACTTATTTTATTTATGATTTTCTAATAAATCCAGTCTTTGTTTT 3045
DB 1927 GTGATATTAGATTTATTTTGTGAGAAATTAATTTTAAATAAATAAATGATTTAT 1986
QY 3046 TTTAAAGAGACTTTAAATATTAAATTTCTTTTAGTGTATTTACCAGTCTTTCCAGGCTA 3105
DB 1987 TTTTGA--ATTTTAAAGATTAATAAATGAGGAAATTTGGGATTTAGTTTATATAAATTT 2043
QY 3106 CTTCTTTTGATTTATTTGGTCTCTATCTTTCTCAAGTTTGTGAATTTGGCTAACTCAT 3165
DB 2044 TTTGTTTGTAGATAGATTTTATTTTGTGTTAGTTAGATGTAATGTTACGATTTTC 2103
QY 3166 TTATCTTTATTTTGTAAATTAGCTCTTTAAATTCATTTCTTTTGATTAACAGCTTCAG 3225
DB 2104 GGTATTATTTGTAATTTTGTGTTTTCGGGTTTAAAGTGAATTTT-----GTTTTAG 2153
QY 3226 TTCTATGCTTTTAAATAAGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTG 3285
DB 2154 TTTTTCGAGTAGTTCGGATTTAGATTTTGTATTTAGTTTGTATTTTGTATTTTGTATTT 2213
QY 3286 AAGTTTGCATGCTTTGAGCAATTAATTTAGGATTAATTTGAAATGTTCTCAGATA-TG 3344
DB 2214 TAGTAGAGATGGGTTTATTTATTTGTTGTTAGTTGTTTAAATTTTGTATTTTAAAGTG 2273
QY 3345 CTTTGTACTTGGCAATTTATCTAAGTTTATGATTTATGATTTATGATTTTCTTTTGGG 3404
DB 2274 ATTTATTCGTTTGTGTTTATTAAGTGTAGATTTATAGCGTGAGTTATTGTTATTTAGTT 2333
QY 3405 CATAAAGCTCTATGCAATTTTGTGCTCTATATTTCTTAAATTTATAAATTTGCTTTA 3464
DB 2334 TATTTTATAAATAATTAATTTTAAATTTAGTGTAGAGAAATTAAGGAAAAATAATTA 2393
QY 3465 AAAAGATTTGCTGCTATTAACATGAATTAAGTCTTTATTTGGACTATA 3513
DB 2394 TTTATTTGTTTATCGTTTGTAGAAATATCGATATTTGATTTGTTTATTA 2442

RESULT 40

US-10-204-708-41
; Sequence 41, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR FILING DATE: 2001-04-06
; PRIOR FILING DATE: 2000-04-06
; PRIOR FILING DATE: 2000-04-06
; PRIOR FILING DATE: 2000-04-07
; PRIOR FILING DATE: 2000-04-07
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98

```

; SEQ ID NO 41
; LENGTH: 8537
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-41

Query Match      2.3%; Score 80.1003; DB 1; Length 8537;
Best Local Similarity 47.1%; Pred. No. 25;
Matches 335; Conservative 0; Mismatches 369; Indels 7; Gaps 3;

QY 2799 TGAACCTTAGTAGTTTATATTACTCGAGAAAATAGTAATTTTCATATGTTATTCAAAATAT 2858
DB 6229 TGTAGGTTTACGTATTTAGAGAGTTGAGAGGAGGAGGATTTAGTTTGGAGGTATAGTTG 6288

QY 2859 TTCATAATGTGGTTAAGATA----ATAAGATTTTCAAATGTATTTTATTCCTTTGATTTT 2914
DB 6289 TAGTGAGTTGTGATTAGGTTATTGTATTTTAAATTTGGGCGATAGAGTAAGATAAAATGAA 6348

QY 2915 TCCTACTATTTTAAATTTTCGGATTTTAACTATTTCTTCAATGACTGTATTTCTAATAT 2974
DB 6349 TATATATTTAATGAATATATTTATTAATAGAAATGGTTTTTAAAGATTTTCTTTGTTAGTAT 6408

QY 2975 TTACTTATTCCTATTACITTAATATGCACCTATTTTTATTTGATTTTCTTAATAAAATCCA 3034
DB 6409 TTAATAATTTTTTTTTTTATAGTGGTAATTTGTTATATCGTTTTTAATCAGATGAATTA 6468

QY 3035 GTCCTGTTTTTTTAAAGACITTTAAATATATTAATTTCTCTTAGTGTTTTACCAGTT 3094
DB 6469 ATCTGTGTTGTTTTTGAATTTTATATAAAATATATTAGGAGTGTTCAGTGTTCGTGTTG 6528

QY 3095 CTTTCAGGCTACTCTTTTGAATTTATTTGGTCCTATCTTTTCTCAAGTTTGAATTCGCT 3154
DB 6529 TTTATATGTTTGTAAACGATATTTTTAGTTTTTTTGATTTTAAATTTATTTTGTGTTGTT 6588

QY 3155 ACGTA--ACTCATTTATCTTTATTTTTTGTAAATAGCTCTTTAAATCATATTTCTTTG 3212
DB 6589 ATATAGTAATAGTTTAAATGTGAATTTTAAATTTTTTTTGATTAATAATGAATGATATTTC 6648

QY 3213 ATAACAGCTTCAGTCTCATGGCTTTAATAAGATTTTTTTTTTTTTTTTTTTTTTTTAAAGAA 3272
DB 6649 TTTATATGTTTTTGTGTTTATTTTGTATATTTTTTGTGATGTGTTGTTTAAATTTTTGTTT 6708

QY 3273 GTCATTCCTTTGTGAAGTTTTCACAATGCTTTGACAAATAATTTAGGATATTTTGAATGG 3332
DB 6709 ATTTTT--TTATATTTTTTTTTTTTATGATGTGAAGTTTTTTTATATATTTGGTTATTAGTTT 6767

QY 3333 TTCATGAGTATGCTTTTGTACTTCGGCATTTATTTGAAGTTTATGATTTTATGAATTTATGATG 3392
DB 6768 TTAGTTATGTTTGTGTTAAATGTTTTTTTTTATTTGTTGTTGATTTGTTATTTTTTTTAGTG 6827

QY 3393 CTTTTTTTTTGGGCATAAAGGCTCATGCGATATTTTTTGTGCTCATATTTCTTAAAAATAT 3452
DB 6828 GTGTTTTTTGATGAATAGAAGTGTGAATTTATTTTTTTTATTAATTTATTTATTTATTTAT 6887

QY 3453 AAATTTGGCTTTAAAGATTTTTCGTCTATTTAAACATGAATTAAGTCTTAT 3503
DB 6888 TTTTTTTTTTCAGATAGAGTTTGTGTTTTTGTGTTTAGTTTGGAGTCTAAT 6938

```

RESULT 41
US-08-998-416-186
; Sequence 186, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Bohmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jürgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebschung, Corinne

```

, TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHEYA GOSSYPII
,
, TITLE OF INVENTION: AND USES THEREOF
,
, NUMBER OF SEQUENCES: 1152
,
, CORRESPONDENCE ADDRESS:
,
, ADDRESSEE: No. 6239264artis Corporation
,
, STREET: 3054 Cornwallis Road
,
, CITY: Research Triangle Park
,
, STATE: No. 6239264th Carolina
,
, COUNTRY: USA
,
, ZIP: 27709
,
, COMPUTER READABLE FORM:
,
, MEDIUM TYPE: Floppy disk
,
, COMPUTER: IBM PC compatible
,
, OPERATING SYSTEM: PC-DOS/MS-DOS
,
, SOFTWARE: PatentIn Release #1.0, Version #1.30
,
, CURRENT APPLICATION DATA:
,
, APPLICATION NUMBER: US/08/998,416
,
, FILING DATE: 24-DEC-1997
,
, CLASSIFICATION: 435
,
, PRIOR APPLICATION DATA:
,
, APPLICATION NUMBER: CH 0016/97
,
, FILING DATE: 31-DEC-1996
,
, ATTORNEY/AGENT INFORMATION:
,
, NAME: Meigs, J. Timothy
,
, REGISTRATION NUMBER: 38,241
,
, REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: 919-541-8587
,
, TELEFAX: 919-541-8689
,
, INFORMATION FOR SEQ ID NO: 186:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 615 base pairs
,
, TYPE: nucleic acid
,
, STRANDEDNESS: single
,
, TOPOLOGY: linear
,
, MOLECULE TYPE: DNA (genomic)
,
, ORIGINAL SOURCE:
,
, ORGANISM: PAG1074RP
,
, US-08-998-416-186

```

Query Match	2.2%; Score 79.7003; DB 1; Length 615;
Best Local Similarity	49.6%; Pred. No. 3.3e+02;
Matches	280; Conservative 0; Mismatches 278; Indels 7; Gaps 3;
QY	2843 ATGPAATCAAAATATTATTCATAATGTTGGTTAAGATAAATAAGATTTTCAAATGATTTTT 2902
DB	22 ATTTATAAAGATTAATAATAAACTTTTTATTATAATATTTAAGTATTAATATTAAAC 81
QY	2903 ATCTTTCATTTTTCTCTACTATTATTAATTTGGGATTTTAACTATTTCTTCCAAGACTTG 2962
DB	82 TATTATTATCATTTATTTAATAAAATAATTTGATTTAATAACTTTATTATATAATTTAT 141
QY	2963 TATTTCTAAATATTTACTATTATTCATTTTCTTAAATGCACCTATTTTTATTGATTTTTTC 3022
DB	142 TATATAATTTACTTAATTCATCATATTATTAATTTATTAATTAATAAAATAATATTAA 201
QY	3023 T---AAATAAAATCCAGTCCTTGTT---TTTTAAAAAGACTTTAAAAATTAATAATTTCTC 3076
DB	202 TATGAATACTATTTAGTCTATGTTCAAAATTTTAAATTAGTTATTTAAAAATATTATTAGATA 261
QY	3077 TTTAGTGTTTACACAGTCTCTTCAGGCTACTCTTTTGATTTATTTGGTCTCATCTTTTC 3136
DB	262 TTATATTTCTTTTAAATAAATATTAATAATAGATTATCAATTAATTAATATTATTATTA 321
QY	3137 TCAAGTTTTGAATTTGGCTAGTCACTCATTTATCTTTATTTTTTGTAAATPAGTCTTTA 3196
DB	322 ATTTGTTATTTAAAAATAATATATTTTATTATATAAAGATTTAAATTTATTTAAATATGTA 381
QY	3197 AATTCATATTTCTTTGATAACAGCTTCAGTCTATAGCTTTTAATAAAGCTTTTTTTTTTT 3256
DB	382 AATTATTATTTTATATAATATCTATTTTATATAAAATATTATGTTGATTATTAATTTTA 441
QY	3257 TTTTTTTTTTAAAGAAATGTCATTTCTTTGTGAAGTTTTTGACAAATGTTTGAGCAATAATTTA 3316

Db 442 ACTTTTATAAGAAATTATTATTAAGATTAAATTTAACTTTAAATTCCTATTATTAAATTT- 500
Qy 3317 GGATATTTTGAATGGGTTCAAGCATGCTTTGTACTTGGCATTATTGAAGTTTATGA 3376
Db 501 TTATATATTATTAATAAATTATATTCATTTTATTATTATTATTAATTAATTAATTA 560
Qy 3377 TTTATGAATTATGATGCTTTTTTTT 3401
Db 561 TTTAATTAAATTTTATCATTATTT 585

RESULT 42
US-10-204-708-71
; Sequence 71, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: ZIEFENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 71
; LENGTH: 8607
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-71

```

Db      6516  AATTTTTTAGTGATTGTGTGTTGTAATAATTAGTGAATAATAATTTTTTTTTTATGT 6575
Qy      3206  TTCTTTGATAACAGCCTTCAGTCTCTATGGCTTTTAATAAAGTTTTTTTTTTTTTTTTTTTTT 3265
Db      6576  GTTCTTTGAGATATAATTATATTATTAGAAATTATTATAATTTTTTTTGATAGGAATTAAT 6635
Qy      3266  AAGAATGTCATCTTTTGGAAGTTTGGACAAATGCTTTGAGCAATAATTTTAGAGATAATTT 3325
Db      6636  TAGGTTTTTAAATAATTTTTAGCGTTAAGAGATTAATTTAAAAAAGAGTAGTTTTTTTTTAT 6695
Qy      3326  TGAATGTTTCATGAGTATGCTTTTGTACTTGGCATTTATTGAAGTTTATGATTTATGAAT 3385
Db      6696  GATTATTTTTTAAATGTTATTTTTTATTTTCGTATGGAGTCGATTAGTTTTTTTCGTAAT 6755
Qy      3386  TATGATGCTTTTTTTTGGGCATAAAGCTCATGGCATATTTTTTTGGTGCATATTTCTTA 3445
Db      6756  TTTTATTTGTGATTCGTGTGTTTTTTTGGTAGATTAATTTTTTTTATATAATAGTTGTT 6815
Qy      3446  AAATTAATAATGGCTTTAAAAGTATTTTGCTGCTATT 3483
Db      6816  AAATATTTGGGGGAGGTTAGATACGGTGGTTTATATT 6853

RESULT 43
US-10-204-708-26
; Sequence 26, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 26
; LENGTH: 6583
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-26

```


; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 39
; LENGTH: 19513
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-39

```
Query Match          2.2%; Score 77.6002; DB 1; Length 19513;
Best Local Similarity 48.2%; Pred. No. 12;
Matches 327; Conservative 0; Mismatches 339; Indels 12; Gaps 4;

QY 2806 TTAGTAGTTTATATACCTCAGAAATAGTAAATTCATATGATGATTCACAAATATTTTCATAA 2865
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 12679 TTATATTTTATTTTAAATTTAAAGTTTAAATTTAGGATATATATATGTTATGTTAA 12738
QY 2866 TGTGGTTAAGATATTAAGATTTTCAAAATGATTTTATCTTTCATTTTCTCTACTTAT 2925
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 12739 TTTTATTATTTTAAAGATAATGTT---TGTGTTTTTGTGTTAGATGTTAGTGGT 12794
QY 2926 TTAATTTTGGGATTTTAACTATTTCTTCAATGACTGTTATTTCTAATATTTACTTATCT 2985
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 12795 GTAGTTATGG--TTTATTTGTTAGTTTATTTTATTTAGGTTTAAAGTATTTTATTTAG 12852
QY 2986 ATTTTACTTTAATGCACTTATTTTATTTATGATTTTCTTAATAAATCCAGTCCTTGTGTT 3045
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 12853 TTTTAAAGTATGTTGGGATTTAGGATGATGATTTATTTTAAAGT--AAATTTTAAATTT 12910
QY 3046 TTTTAAAGACCTTTAAATTTATTTCTTTTCTTTAGTGTTTTACCAGTCTTTTCAGGCTA 3105
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 12911 TTGTAGAGATGATTTAAATTTTAAAT---TTTGTAGAAATGGAGTTTATTTATGTTG 12966
QY 3106 CTCTTTTGAATTTTGGTCTATCTTTCTCAAGTTTGTGAATTTGGCTACGTAACATCAT 3165
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 12967 TTTAGTTTATGTTTAAATTTTGGTTTAAAGTGATTTATTTTATTTTAAAGTG 13026
QY 3166 TTATCTTTATTTTGTAAATAGCTCTTTAAATTCATTTCTTTGATTAACAGCTTCAG 3225
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13027 TTGGATTTATAGACGTGAGTTTTTTATGTTAGGTAATTTATTTTATGATATCGATAG 13086
QY 3226 TTCTATGGCTTTAATAAAGTTTATTTTATTTTATTTTAAAGAAATGTCATCTTTGTG 3285
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13087 TTTTAAAGATTTATAGATTAGTTGTTATTTTATTTTGGATTCGTTGTTTTTTTTT 13146
QY 3286 AAGTTTGCACAACTTTTGACAAATAATTTAGGAATTTTGTGAATGTTTCATGAGTATGC 3345
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13147 TTGTTTGTGTTTGAGATAGAGTTTGTGTTGTTATTTAGGTTGGAGTGTAGTGGCGTAT 13206
QY 3346 TTTTGTACTTGGCATTTTATGAGTTTATGATTTTGAATTTATGATCTTTTGTGGGC 3405
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13207 CGTAGTTTATGTAATTTTGTGTTTATTTAGGTTTAAAGGATTTTGTGTTTTAGTTT 13266
QY 3406 ATAAAGCTCATGGCATATTTTGTGCTCTATATTTCTTAAATTTATTAATTTGGCTTTAA 3465
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 13267 AGTAGTTGGTATTTAGGTAATTTATTTATTTAGCTTTGGTTAAATTTTGTATTTTAGTAGAG 13326
QY 3466 AAAGTATTTGCTGCTATT 3483
DB ||||| ||||| |||||
DB 13327 ATGGGGTTTATTTATGTT 13344
```

Search completed: August 24, 2004, 13:31:55
Job time : 614 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 13:34:23 ; Search time 279 Seconds
(without alignments)
3.941 Million cell updates/sec

Title: US-10-664-775-2

Perfect score: 3572
Sequence: 1 gtcaggaaggcgagctga.....gcaacacagcagaagctt 3572

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 0.5

Searched: 43 seqs, 133908 residues

Total number of hits satisfying chosen parameters: 86

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rnpb2.seq *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3572	100.0	3572	1	US-10-664-775-2
2	1012.1	28.3	78056	1	Sequence 2, Appli
3	966.804	27.1	23651	1	Sequence 1, Appli
C 4	885.401	24.8	78056	1	Sequence 2, Appli
5	557.9	15.6	23651	1	Sequence 1, Appli
C 6	413.4	11.6	908	1	Sequence 2, Appli
7	380.6	10.7	967	1	Sequence 1, Appli
8	372.6	10.4	419	1	Sequence 13832, A
9	364.7	10.2	451	1	Sequence 844, App
10	363.2	10.2	436	1	Sequence 6161, Ap
11	360.3	10.1	432	1	Sequence 5090, Ap
12	357.1	10.0	433	1	Sequence 4684, Ap
13	353.3	9.9	433	1	Sequence 5347, Ap
14	347.4	9.7	436	1	Sequence 7158, Ap
15	344.199	9.6	31412	1	Sequence 3, Appli
16	335.8	9.4	428	1	Sequence 8154, Ap
C 17	333	9.3	406	1	Sequence 1308, Ap
C 18	324	9.1	423	1	Sequence 8311, Ap
19	321.6	9.0	410	1	Sequence 13834, A
C 20	320.8	9.0	423	1	Sequence 5499, Ap
C 21	320.4	9.0	409	1	Sequence 14308, A
22	313.4	8.8	425	1	Sequence 9545, Ap
23	307.8	8.6	422	1	Sequence 3391, Ap
24	299.8	8.4	336	1	Sequence 11214, A
25	299.4	8.4	447	1	Sequence 11945, A
C 26	296	8.3	410	1	Sequence 4133, Ap
C 27	289.4	8.1	356	1	Sequence 3113, Ap
28	286	8.0	413	1	Sequence 13876, A
29	285.2	8.0	405	1	Sequence 9592, Ap
C 30	284.5	8.0	407	1	Sequence 13585, A
31	282.9	7.9	423	1	Sequence 13323, A
32	279.1	7.8	365	1	Sequence 13570, A
33	275.7	7.7	448	1	Sequence 10238, A

ALIGNMENTS

RESULT 1

US-10-664-775-2
; Sequence 2, Application US/10664775
; Publication No. US2004011576A1
; GENERAL INFORMATION:
; APPLICANT: Simesen, Ruth B
; APPLICANT: Pedersen, Anette A
; APPLICANT: Faisst, Steffen
; APPLICANT: Jensen, Jan J
; APPLICANT: Weiguny, Dietmar
; TITLE OF INVENTION: Method for Making Recombinant Proteins
; FILE REFERENCE: 6449.200-US
; CURRENT APPLICATION NUMBER: US/10/664,775
; PRIOR FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: Danish Application PA 2002 01384
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US 60/416,566
; PRIOR FILING DATE: 2002-10-07
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 3572
; TYPE: DNA
; ORGANISM: Baby hamster kidney cell line
US-10-664-775-2

Query Match	100.0%;	Score	3572;	DB 1;	Length	3572;			
Best Local Similarity	100.0%;	Pred. No.	1.4e-13;	Mismatches	0;	Indels	0;	Gaps	0;
Matches	3572;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;

Qy	1	GTGAGGAGGGGGCGGCGAGTGTGAGGAGGTTACTTACCTCGTCCAGGTAGGAGGAGTACG	60
Db	1	GTGAGGAGGGGGCGGCGAGTGTGAGGAGGTTACTTACCTCGTCCAGGTAGGAGGAGTACG	60
Qy	61	TGGCGTTTGTGTGAGGAGCGGCGTAAAGAGATACCCACGCCCGCCAGGTAAGAGAAACCCCAAG	120
Db	61	TGGCGTTTGTGTGAGGAGCGGCGTAAAGAGATACCCACGCCCGCCAGGTAAGAGAAACCCCAAG	120
Qy	121	TAAGATGTAGTGTGTGTGAGAGGGGCGATCAGAGGGGCGACATCTGTAACCAACCATACACGCA	180
Db	121	TAAGATGTAGTGTGTGTGAGAGGGGCGATCAGAGGGGCGACATCTGTAACCAACCATACACGCA	180
Qy	181	GAAACTAGTCAATCTAATCACTAGGACACAGCGTTGTCTAATCAATGAACTAAG	240
Db	181	GAAACTAGTCAATCTAATCACTAGGACACAGCGTTGTCTAATCAATGAACTAAG	240
Qy	241	CCATGCCCGTGGGGCAACCAAGATGGGAGGTCATGTGGGAGAGATCTGACAGAAATGTG	300
Db	241	CCATGCCCGTGGGGCAACCAAGATGGGAGGTCATGTGGGAGAGATCTGACAGAAATGTG	300
Qy	301	GTCCACTGGAGAGGAATGCAACCACTTCAGTATTTCTTGCCTTCAGAACCCCATGAAC	360
Db	301	GTCCACTGGAGAGGAATGCAACCACTTCAGTATTTCTTGCCTTCAGAACCCCATGAAC	360
Qy	361	AGTATCAAAAGGCAAAATGATAGTACTGAAAGAGAGAACTCCCGAGGTCAAGTGTGCC	420
Db	361	AGTATCAAAAGGCAAAATGATAGTACTGAAAGAGAGAACTCCCGAGGTCAAGTGTGCC	420

Db 361 AGTATGAAAAGCCAAATGATAGGATACCTGAAAGAGGAACTCCCCAGGTCACTAGTGGTCC 420
Qy 421 CCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCCA 480
Db 421 CCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAAATGAAGAGATGGAGCCA 480
Qy 481 AAGCAAAAAGAAATACCCAGCTGTGGATGTGATCTGGTGTATTAAGCAAGTCCGATGCTGT 540
Db 481 AAGCAAAAAGAAATACCCAGCTGTGGATGTGATCTGGTGTATTAAGCAAGTCCGATGCTGT 540
Qy 541 AAAGAGCAATATGTCATAGGAACCTGGAAATGTCAGGTCCATGAATCAAGGCAAAATGGAA 600
Db 541 AAAGAGCAATATGTCATAGGAACCTGGAAATGTCAGGTCCATGAATCAAGGCAAAATGGAA 600
Qy 601 GTGGTCAAAACAGAGATGGCAAGTGTGAATGTCAACATTTAGGAATCAGCGAACTAAAA 660
Db 601 GTGGTCAAAACAGAGATGGCAAGTGTGAATGTCAACATTTAGGAATCAGCGAACTAAAA 660
Qy 661 TGGACTGGAATGGGTGAATTTAACTCAGATGACCATATATCTACTACTGCGGGCAGGAA 720
Db 661 TGGACTGGAATGGGTGAATTTAACTCAGATGACCATATATCTACTACTGCGGGCAGGAA 720
Qy 721 TCCCTCAGAAAGAAATGGAGTAGCCATCATGTGTCAACAAAGAGTCCGAAATGCGACTTT 780
Db 721 TCCCTCAGAAAGAAATGGAGTAGCCATCATGTGTCAACAAAGAGTCCGAAATGCGACTTT 780
Qy 781 GGATGCGACTCAAAACAGCAGAGATGATCTCTGTTTGTTCAGGCAAAACCAATTCAT 840
Db 781 GGATGCGACTCAAAACAGCAGAGATGATCTCTGTTTGTTCAGGCAAAACCAATTCAT 840
Qy 841 ATCAGTAATCCAGTCTATGCCCAACAGTAATGCTGAAGAGCTGAAGTTGAACGG 900
Db 841 ATCAGTAATCCAGTCTATGCCCAACAGTAATGCTGAAGAGCTGAAGTTGAACGG 900
Qy 901 TCCTATGAAGACCTCAAGACCTTTTGAAGTAAACACCCAAAGAGTCTCTTCTCAT 960
Db 901 TCCTATGAAGACCTCAAGACCTTTTGAAGTAAACACCCAAAGAGTCTCTTCTCAT 960
Qy 961 ATAGGGAGTCAATGCAAAAGTAGGAAGCAAGAAACACCTGAGTAAACAGCAAAATTT 1020
Db 961 ATAGGGAGTCAATGCAAAAGTAGGAAGCAAGAAACACCTGAGTAAACAGCAAAATTT 1020
Qy 1021 GGCCTTGGAAATCGGAATGAAGCAGGCAAGACTAATAGAGTTTGGCCAAAGAAATGCA 1080
Db 1021 GGCCTTGGAAATCGGAATGAAGCAGGCAAGACTAATAGAGTTTGGCCAAAGAAATGCA 1080
Qy 1081 CTGGTCAAGCAACCCCTCTCCACACACAGAGAGAGCTCTACATGGACATCA 1140
Db 1081 CTGGTCAAGCAACCCCTCTCCACACACAGAGAGAGCTCTACATGGACATCA 1140
Qy 1141 CCAGATGGTCAACACCGAAATCAGATGATTAATTTCTTTCAGCCCAAGAGATGGAGAGC 1200
Db 1141 CCAGATGGTCAACACCGAAATCAGATGATTAATTTCTTTCAGCCCAAGAGATGGAGAGC 1200
Qy 1201 TCTATACAGTCAAGCAAAACAGACAGGAGCTTACTGGCTCAGATCATGAACTCCTT 1260
Db 1201 TCTATACAGTCAAGCAAAACAGACAGGAGCTTACTGGCTCAGATCATGAACTCCTT 1260
Qy 1261 ATTGCCAAATTCAGACTTAAATTAAGAAAGTAGGAAACCACTAGATCACTCAGGTAA 1320
Db 1261 ATTGCCAAATTCAGACTTAAATTAAGAAAGTAGGAAACCACTAGATCACTCAGGTAA 1320
Qy 1321 GACCTAAATCCATCCCTTATGATTAATACAGTGAAGTGAAGAAATAGATTAAGGCGCTA 1380
Db 1321 GACCTAAATCCATCCCTTATGATTAATACAGTGAAGTGAAGAAATAGATTAAGGCGCTA 1380
Qy 1381 GATCTGATAGCAGAGTACCTTAATGAATGATGACAGAGGTTTCACTGACATTTGACAGGAG 1440
Db 1381 GATCTGATAGCAGAGTACCTTAATGAATGATGACAGAGGTTTCACTGACATTTGACAGGAG 1440
Qy 1441 ACAGGATCGAGACCATCCCATGGAAAGAAATGCAAAAGCAAAATGGCTGTCTGG 1500
Db 1441 ACAGGATCGAGACCATCCCATGGAAAGAAATGCAAAAGCAAAATGGCTGTCTGG 1500

Qy 1501 GAGGCCCTTCAAAATAGCTGTGAAAAAGAAAGAGAGTGAAGGCAAGAAAAAGGAAAGA 1560
Db 1501 GAGGCCCTTCAAAATAGCTGTGAAAAAGAAAGAGAGTGAAGGCAAGAAAAAGGAAAGA 1560
Qy 1561 TAAAGCATCTGAATGAGAGTTCMAAGAACTTCAGTTGTTCAAGCTGGTTTAGAAA 1620
Db 1561 TAAAGCATCTGAATGAGAGTTCMAAGAACTTCAGTTGTTCAAGCTGGTTTAGAAA 1620
Qy 1621 AGTCAGAGGAAACAGAGACCAATTTGCCAACTCTCTGTATCATGGAAGAAAGCAAGAGA 1680
Db 1621 AGTCAGAGGAAACAGAGACCAATTTGCCAACTCTCTGTATCATGGAAGAAAGCAAGAGA 1680
Qy 1681 GTTCCAGAAAAACATCTATTTCTGCTTTATTCAGTATGCAAAAGCTTTGACTGTGGGG 1740
Db 1681 GTTCCAGAAAAACATCTATTTCTGCTTTATTCAGTATGCAAAAGCTTTGACTGTGGGG 1740
Qy 1741 TCACATTAACCTGTGAAAAATTTCTGAAAGGATGGAAATACCAGACCACTGACCTGACT 1800
Db 1741 TCACATTAACCTGTGAAAAATTTCTGAAAGGATGGAAATACCAGACCACTGACCTGACT 1800
Qy 1801 CTTGAAAAATTTGTATGTCAGGTTCAGAGAACACAGTTAGAACTGGACATGGAACACAGA 1860
Db 1801 CTTGAAAAATTTGTATGTCAGGTTCAGAGAACACAGTTAGAACTGGACATGGAACACAGA 1860
Qy 1861 CTGGTTCAGTAGGAAAGGAGTATGCAAGCTGTATATTTGTCACCGGCTTGTTTAA 1920
Db 1861 CTGGTTCAGTAGGAAAGGAGTATGCAAGCTGTATATTTGTCACCGGCTTGTTTAA 1920
Qy 1921 CTTCTATGCAAGACATCATGAGAAACCTGGGCTGGAAAGCAACAGCTGGAAATCAAG 1980
Db 1921 CTTCTATGCAAGACATCATGAGAAACCTGGGCTGGAAAGCAACAGCTGGAAATCAAG 1980
Qy 1981 ATTGCGGGAGAAATAGCAATTAACCTCAGATATGCAAGTATGCAACCTTATGCGAGAA 2040
Db 1981 ATTGCGGGAGAAATAGCAATTAACCTCAGATATGCAAGTATGCAACCTTATGCGAGAA 2040
Qy 2041 AGTGAAGAGGAACTAAAAAGCTCTTGAAGTGAAGAGGAGAGTGAAGAAAGTTGGC 2100
Db 2041 AGTGAAGAGGAACTAAAAAGCTCTTGAAGTGAAGAGGAGAGTGAAGAAAGTTGGC 2100
Qy 2101 TTAAAGCTCAACATTCAGAAAAACAGAGATCATGGCATCTGGTCCATCACTTCATGGAA 2160
Db 2101 TTAAAGCTCAACATTCAGAAAAACAGAGATCATGGCATCTGGTCCATCACTTCATGGAA 2160
Qy 2161 ATAGATGGGAAACAGTGGAAACAGTGTGAGTGAAGAGAGTGAAGAAAGTTGGC 2220
Db 2161 ATAGATGGGAAACAGTGGAAACAGTGTGAGTGAAGAGAGTGAAGAAAGTTGGC 2220
Qy 2221 CTGAGAGTGGTGAATGAGGAACTTAAAGACACTTACTCTCTGGAAAGAAAGTTA 2280
Db 2221 CTGAGAGTGGTGAATGAGGAACTTAAAGACACTTACTCTCTGGAAAGAAAGTTA 2280
Qy 2281 ACCACCTAGATAGCATATTCAGAAAGCAGACATTAACCTTGGCCAAAGAGCCCATCTA 2340
Db 2281 ACCACCTAGATAGCATATTCAGAAAGCAGACATTAACCTTGGCCAAAGAGCCCATCTA 2340
Qy 2341 GTCAAGGCTATGGTCTTTCAGTGGTCAATGATGATGATGAGTGGAGTGGAGTGTGAAGAA 2400
Db 2341 GTCAAGGCTATGGTCTTTCAGTGGTCAATGATGATGATGAGTGGAGTGGAGTGTGAAGAA 2400
Qy 2401 GCTGAGCATGAAGAAATGATGCTTTTGAAGTGGTGGAGAGAGTCTTGGAGTCT 2460
Db 2401 GCTGAGCATGAAGAAATGATGCTTTTGAAGTGGTGGAGAGAGTCTTGGAGTCT 2460
Qy 2461 CTTTGGACTGCAAGAGATCCAAACAGTCCATTCTGAAGGAGATCAGCTTGGGATTTCT 2520
Db 2461 CTTTGGACTGCAAGAGATCCAAACAGTCCATTCTGAAGGAGATCAGCTTGGGATTTCT 2520
Qy 2521 TTGAAAGAAATGATGCTTAAAGCTGAAACTCCAGTACTTTTGGCCACCTGATCAGAAAGCT 2580
Db 2521 TTGAAAGAAATGATGCTTAAAGCTGAAACTCCAGTACTTTTGGCCACCTGATCAGAAAGCT 2580

QY 2581 GACTCAGTGAAGACCTGATGCTGGAGGGATTTGGGGCAGGAGAGAGGGGACGA 2640
 Db GACTCAGTGAAGACCTGATGCTGGAGGGATTTGGGGCAGGAGAGAGGGGACGA 2640
 QY 2641 CAGAGATGAGATGCTGATGGGATCACTGATGAGAGGATTTGGGGTGAAGTTC 2700
 Db CAGAGATGAGATGCTGATGGGATCACTGATGAGAGGATTTGGGGTGAAGTTC 2700
 QY 2701 CTGGAGTTGGTGAAGCAGGAGGAGGCTGCTGCTGAGGATTTGAGGATTTT 2760
 Db CTGGAGTTGGTGAAGCAGGAGGAGGCTGCTGCTGAGGATTTGAGGATTTT 2760
 QY 2761 TGGACACGACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTG 2820
 Db TGGACACGACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTGAGCACTG 2820
 QY 2821 ACTCAGAAATAGTAATTCATATGATATTCAGAAATTAATTCATATGATGATTA 2880
 Db ACTCAGAAATAGTAATTCATATGATATTCAGAAATTAATTCATATGATGATTA 2880
 QY 2881 TAAGATTTTCAATGATTTTATCTTTGATTTTCTGATTTTCTGATTTTCTGAT 2940
 Db TAAGATTTTCAATGATTTTATCTTTGATTTTCTGATTTTCTGATTTTCTGAT 2940
 QY 2941 TAACATTTTCAATGATTTTATCTTTGATTTTCTGATTTTCTGATTTTCTGAT 3000
 Db TAACATTTTCAATGATTTTATCTTTGATTTTCTGATTTTCTGATTTTCTGAT 3000
 QY 3001 CACTATTTTCAATGATTTTATCTTTGATTTTCTGATTTTCTGATTTTCTGATTT 3060
 Db CACTATTTTCAATGATTTTATCTTTGATTTTCTGATTTTCTGATTTTCTGATTT 3060
 QY 3061 AAATATTAAATTTCTTTAGTGTTTTACCAGTCTTTTACAGTCTTTTACAGTCT 3120
 Db AAATATTAAATTTCTTTAGTGTTTTACCAGTCTTTTACAGTCTTTTACAGTCT 3120
 QY 3121 TTGGTCCATCTTTTCAAGTTTGAATTTGGTACGTAAGTCAATTTATCTTTATTTT 3180
 Db TTGGTCCATCTTTTCAAGTTTGAATTTGGTACGTAAGTCAATTTATCTTTATTTT 3180
 QY 3181 GTAAATTTAGTCTTTTAAATTTCAATTTCTTTGATTAACAGCTTTCAAGTTCTA 3240
 Db GTAAATTTAGTCTTTTAAATTTCAATTTCTTTGATTAACAGCTTTCAAGTTCTA 3240
 QY 3241 AAAGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 3300
 Db AAAGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 3300
 QY 3301 TTTGAGCAATAATTTAGGATATTTTGAATGTTTCAATGATGCTTTTGTACTTGGCAT 3360
 Db TTTGAGCAATAATTTAGGATATTTTGAATGTTTCAATGATGCTTTTGTACTTGGCAT 3360
 QY 3361 TTTGAGCAATAATTTAGGATATTTTGAATGTTTCAATGATGCTTTTGTACTTGGCAT 3420
 Db TTTGAGCAATAATTTAGGATATTTTGAATGTTTCAATGATGCTTTTGTACTTGGCAT 3420
 QY 3421 ATATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 3480
 Db ATATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 3480
 QY 3481 ATTAACATGAATTAAGTCTTTATTTGATGATGATGATGATGATGATGATGATGAT 3540
 Db ATTAACATGAATTAAGTCTTTATTTGATGATGATGATGATGATGATGATGATGAT 3540
 QY 3541 CTTAGCGACTAAGCAACACAGCAGCAAGGCTT 3572
 Db CTTAGCGACTAAGCAACACAGCAGCAAGGCTT 3572

RESULT 2

US-10-109-551-1

; Sequence 1, Application US/10109551

; Publication No. US20020194635A1

GENERAL INFORMATION:

APPLICANT: DUNNE, PATRICK W.
 APPLICANT: PIEDRAHITA, JORGE
 TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
 FILE REFERENCE: SPONGIFORM ENCEPHALOPATHIES
 CURRENT APPLICATION NUMBER: US/10/109,551
 CURRENT FILING DATE: 2002-03-28
 PRIOR APPLICATION NUMBER: 60/280,549
 PRIOR FILING DATE: 2001-03-30
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1
 LENGTH: 78056
 TYPE: DNA
 ORGANISM: Bos taurus
 US-10-109-551-1

Query Match 28.3%; Score 1012.1; DB 1; Length 78056;
 Best Local Similarity 92.8%; Pred. No. 0,0026;
 Matches 1113; Conservative 0; Mismatches 79; Indels 7; Gaps 5;

QY 1589 GAACCTTCAGTTGTTCAAGCTGTTTTHAGAAAGTCAGAGAACCCAGAGACCAATTTGCC 1648
 Db GAACCTTCAGTTGTTTCAAGCTGTTTTHAGAAAGTCAGAGAACCCAGAGACCAATTTGCC 33332
 QY 1649 AACATCTCTGATCATGCAAAAAGCAAGAGAGTTCCAGAAAAACATCTATTTCTGCTTT 1708
 Db AACATCTCTGATCATGCAAAAAGCAAGAGAGTTCCAGAAAAACATCTATTTCTGCTTT 33392
 QY 1709 ATGATCTATGCAAAAGCTTTGACCTGTGGGGTCAATAAACTGTGGAAAAATTCGAAA 1768
 Db ATGATCTATGCAAAAGCTTTGACCTGTGGGGTCAATAAACTGTGGAAAAATTCGAAA 33452
 QY 1769 GGGATGGGAATACACAGACCACTGACCTGACTCTTGAAAAATTTGTATGCAAGTCAGAA 1828
 Db GGGATGGGAATACACAGACCACTGACCTGACTCTTGAAAAATTTGTATGCAAGTCAGAA 33512
 QY 1829 GCAACAGTTAGAACCTGAGCATGGAACCAACAGACTGGTTCAGTAGGAAAAAGAGTAGT 1888
 Db GCAACAGTTAGAACCTGAGCATGGAACCAACAGACTGGTTCAGTAGGAAAAAGAGTAGT 33572
 QY 1889 CAAGCTGTATATTCTCACCCTGCTTTTAACTTCTATGTCAGAG-ACATCATCAGAAAC 1947
 Db CAAGCTGTATATTCTCACCCTGCTTTTAACTTCTATGTCAGAG-ACATCATCAGAAAC 33632
 QY 1948 GCTGGCTGGAAGAGCAACAGCTGGAATCAAGATTCGCGGAGAAATAGCAATAAACCCTC 2007
 Db GCTGGCTGGAAGAGCAACAGCTGGAATCAAGATTCGCGGAGAAATAGCAATAAACCCTC 33692
 QY 2008 AGATATGAGATGATACCAACCTTTATGCGAAGAGTGAAGAGGAACTTAAAGGCTCTTG 2067
 Db AGATATGAGATGATACCAACCTTTATGCGAAGAGTGAAGAGGAACTTAAAGGCTCTTG 33752
 QY 2068 ATGAAGTGAAGAGGAGAGTCAAAAAGTTGGCTTTAAAGCTCAACATTCAGAAAAACGAAG 2127
 Db ATGAAGTGAAGAGGAGAGTCAAAAAGTTGGCTTTAAAGCTCAACATTCAGAAAAACGAAG 33812
 QY 2128 ATCATGGCATCTGCTCCCATCATCTGGAATAGTGGGAAACAGTGGAAACAGTG 2187
 Db ATCATGGCATCTGCTCCCATCATCTGGAATAGTGGGAAACAGTGGAAACAGTG 33872
 QY 2188 TCAGACTTTATTTTGGGGGCTCCAAAATCACTGCGAGATGTTGCTGACGCAATGAAA 2247
 Db TCAGACTTTATTTTGGGGGCTCCAAAATCACTGCGAGATGTTGCTGACGCAATGAAA 33930
 QY 2248 TTAAGAGACCTTACTCTCTGGAGAAAGTTA-ACCACTAGATAGCATATTGAAA 2305
 Db TTAAGAGACCTTACTCTCTGGAGAAAGTTA-ACCACTAGATAGCATATTGAAA 33990
 QY 2306 GCAGAGACATTACCTTGGCAACAGCCCTCTAGTCAAGGCTATGTTTCCAGTGG 2365
 Db GCAGAGACATTACCTTGGCAACAGCCCTCTAGTCAAGGCTATG-TTTTCTCTGG 34048

QY 2833 GTAATTTCAATGATTAATCAAAATATTTCATAAATGTTGGTTAGATAAATAAGATTTCAA 2892
Db 12108 GTGTCATGATTAATTAAGGACCTTTAAATTTGCAATTTTCTGA-----AGATTGAT 12058
QY 2893 ATTGATTTTATCTCTGATTTTCTCTACTATTATTAATTTGGCAATTTAACTATTTCIT 2952
Db 12057 GATGTAGAGATTTTTCATGTTTATTAATGCAATTAAGATAACTTCTTTTGCAAGTTCC 11998
QY 2953 CAATGACTTCTATTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3012
Db 11997 CATTCAGGTCTATTGTCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 11938
QY 3013 TT 3014
Db 11937 TT 11936

RESULT 4
US-10-109-551-1/c
; Sequence 1, Application US/10109551
; Publication No. US20020194635A1
; GENERAL INFORMATION:
; APPLICANT: PIEDRAHITA, JORGE
; APPLICANT: PIEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; TITLE OF INVENTION: SPONGIFORM ENCEPHALOPATHIES
; FILE REFERENCE: TAMK:20705
; CURRENT APPLICATION NUMBER: US/10/109,551
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,549
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 78056
; TYPE: DNA
; ORGANISM: Bos taurus
US-10-109-551-1

Query Match 24.8%; Score 885.401; DB 1; Length 78056;
Best Local Similarity 86.6%; Pred. No. 0.0093;
Matches 1063; Conservative 0; Mismatches 151; Indels 14; Gaps 8;

QY 1573 AATGACAGAGTTCCAAAGAACTTCCAGTTGTTTCAAGCTGGTTTAAAGAGTTCAGAGAAC 1632
Db 1660 AAGTACATGACCATGAACTTCCAGATGTTTCAAGCTGGTTTAAAGAGTTCAGAGAAC 1601
QY 1633 CAGAGACCAAAATGCGACATCTCTGATCATGGAAGAAAGCAGAGAGTTCCAGAAAAA 1692
Db 1600 CAGAGATCAAAATGCGACATCTCTGATCATGGAAGAAAGTAAAGAGTTCAGAGAAAA 1541
QY 1693 CATCTATTCTCTTTTATTGACTATGCAAAAGCCTTTCAGCTGTGGGGTCAATAAACT 1752
Db 1540 CATCAATTTTCTTTTATTGACTGTGCAAGCCTTTCAGCTGTGGGATCACAACT 1481
QY 1753 GTGGAAAATTTCTGAAAGGATCGGAATACCAAGACCTGACCTGACTCTTTGAAAAATTT 1812
Db 1480 GTGGGAAATTTCTTAAAGAGATGGAATACCAAGAGCCTGACCTGCTTCCCTGAGAAATCT 1421
QY 1813 GTATGAGTTCAGAGACCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1872
Db 1420 GTATGCAAGTTAGGAAGCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 1361
QY 1873 AGGAAAAGGATGATGTCAGAGGCTGTATATTGTCACCCGGCTTGTGTTAACTTCTATGCAGA 1932
Db 1360 CAGGAAGAGGATGATGTTGAGGCTGTATATTGTCACCCCTGCTTATTTAATTTATATGCAGA 1301
QY 1933 G-ACATCATGAGAAAGCTGGGCTGGGAGAGCAAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 1991
Db 1300 GTACATCAGCAAAATGCTGGAGTGGATTAAGCAAGCTGGAATTAAGTTTGGCCAGGAG 1241
QY 1992 AATATCAATTAACCTCAGATATCAGATGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2051

Db 1240 AATATCAATTAACCTCAGATATCAGATGATGACCACTCCTTATGG--TTAAGAGAGAGA 1183
QY 2052 ACTAAAAAGCCCTCTGTGATGAGGTGAAGAGAGAGAGTGAAGAGTGTGCTTTAAAGCTCAA 2111
Db 1182 CTTAAAGAGCCCTTTTGTGATGAAGTGAAGAGAGAGAGTGAAGAGTGTGCTTTAAAGCTCAA 1123
QY 2112 CATTCAGAAAAAGAGATCATGAGTCTGTTCCCATCTCTTCAATGGAAGATAGATGGGGA 2171
Db 1122 CATTCAGAAAAAGATCATGAGTCTGTTCCCATCTCTTCAATGGAAGATAGATGGGGA 1063
QY 2172 AACAGTGAAGACAGTGTGAGACTTTTATTTTGGGGGCTCCAAATCACTGCGAGATGCT 2231
Db 1062 AACAGTGAAGACAGTGTGAGACTTTTATTTTGGGGGCTCCAAATCACTGCGATATGCT 1006
QY 2232 GACTGAGCCATGAAATTAAGACACTTACTCTTGGAGAAAGATTTA--ACCAACCTA 2289
Db 1005 GACTGAGCCATGAAATTAAGAGTCTTCTCTTGGAGAAAGATTTATGACCAACCTA 947
QY 2290 GATAGCATATTGAAAGAGAGACATTAATCTTCCCATGAGAGTTCAGAGTCTGAGTCT 2349
Db 946 GACAGCATATTAAAAAGAGAGACATTAATCTTGGCCAAAGAGTCCATCTAGTCAAGCT 887
QY 2350 ATGGTTTTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2409
Db 886 ATGGTTTTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 827
QY 2410 TGAAGAAATTTGATGCTTTTGAATCTGTTGTTTGGAGAGACTCTTGGAGTCTTGGACT 2469
Db 826 TGAAGAAATTTGATGCTTTTGAATCTGTTGTTTGGAGAGACTCTTGGAGTCTTGGACT 767
QY 2470 GMAAGAGATCAACCAAGTCCATTTCTGAAGAGATGAGAGTGGGATTTCTTTTGGAGAG 2529
Db 766 GCA---AGATCAACCAAGTCCATTTCTGAAGAGAAATGAGTCTTGAATTTATTCGGAAGGA 710
QY 2530 ATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACTCTGATCAGAGAGAGTCTGACTCAGTG 2589
Db 709 CTATGCTGAAGCTGAAGTCCATATCTTTGGCCACTCTGATGCAAGAACTGCTCTCATTTG 650
QY 2590 GAAAAACCTGATGCTGAGAGGATTTGGGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2649
Db 649 GAAAAACCTGATGCTGAGAGATTTGAAGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 590
QY 2650 AGATGCTGATGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2709
Db 589 AGATGCTGATGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 530
QY 2710 GTGATGAGAC--AGGAGAGCCCTGCTCTCGGGGATTCATGAGGTCACAAAGAGTTGGACAG 2768
Db 529 GTGATGAGACAAAGGAGAGCCCTGCTGATGCTGAGTCCATGAGGTTGCAAGAGAG--TCGATACG 471
QY 2769 ACTGAGCACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 2796
Db 470 ACTGAGCACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACTGAACT 443

RESULT 5
US-09-884-586A-2
; Sequence 2, Application US/09884586A
; Publication No. US20030046716A1
; GENERAL INFORMATION:
; APPLICANT: Echeland, Yann
; APPLICANT: Meade, Harry M.
; APPLICANT: Eichner, Wolfram
; APPLICANT: Sommerweyer, Klaus
; TITLE OF INVENTION: TRANSGENICALLY PRODUCED PLATELET DERIVED
; TITLE OF INVENTION: GROWTH FACTOR
; FILE REFERENCE: 10275-120001
; CURRENT APPLICATION NUMBER: US/09/884,586A
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: 60/212,406
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 2
; LENGTH: 23651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(23651)
; OTHER INFORMATION: n = A, T, C or G
US-09-884-586A-2

Query Match      15.6%; Score 557.9; DB 1; Length 23651;
Best Local Similarity 78.4%; Pred. No. 0.52;
Matches 820; Conservative 0; Mismatches 191; Indels 35; Gaps 13;

QY 1744 CAATAAAGTGTGAAAATTCTGAAGGGATGGGAATACAGACCACCTGACCTGACTCTT 1803
DB 2456 CAAACAAACTGAAAATTCTTCAAGAGA--AGAAATACAGACCACCTGCTTCTCT 2513
QY 1804 GAAAAATTTGTATGCAGGTCAGGAAGCAACAGTTAGAACTGGACATGGAACAGAGCTG 1863
DB 2514 GAGAAATCTGTTGTCTCA-GAAGCAACAGTTAGAACCAAGATGGAACAGAGCTG 2572
QY 1864 GTTCAAGTAGGAAAAGAGTAGTGTCAAGGGCTGTATATTGTCACCCGGCTTTGTTAACTT 1923
DB 2573 GTTCCAAATCAGAAAGAGTAGTGTCAAGGGCTGTATATTGTCACCCCTGATTATTAACTT 2632
QY 1924 CTATCAGAG-ACATCATGAGAAAGCTGGCTGGAGAACACAGCTGAATCAAGAT 1982
DB 2633 ATATGCAATGATACATAAATAAC-CAGATACAAGAT-ACACACACTTATGGCAGAAA 2750
QY 1983 TGCCGGGAAATAGCAATAAATCAGATATGCAGATATACCACTTATGGCAGAAAG 2042
DB 2693 TTCTGGGAAATATCAATAAAC-CAGATACAAGAT-ACACACACTTATGGCAGAAA 2750
QY 2043 TGAAGAGAACTAAAGAGCTTGTGATGAAGTGGAAGAGAGAGTGAAGAGTTGGCTT 2102
DB 2751 CTAGAAGAACTAAAGAGCTTGTGATGAAGTGGAAGAGAGAGTGAAGAGTTGGCTT 2810
QY 2103 AAAGTCAACATTCAGAAAACGAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAT 2162
DB 2811 AAAACCCAACTTCAAAATCAGA-----TCAATATTTCATGGCAAT 2853
QY 2163 AGATGGGAAACAGTGGAAACAGTGTGACAGCTTTATTTTGGGGGCTCCAAATCACT 2222
DB 2854 AAATGGGAAACAAATGGAAACAGTGGAGAGCTTTATTTTCT--TGCGCTCCAAATCACT 2911
QY 2223 GCAGATGTGACTGCAGCATGAATTAAGACACTTACTCTCTTGAAGAAAGTTA-- 2280
DB 2912 GCAGATGTGACTGCAGCATG--ATTAAAGATGCTTGTCTCTTGGAGAGAGCTATT 2969
QY 2281 ACCAACCTTAGATAGCATATTGAAAGCAGAGACATTAACCTTGGCCAAACAAAGCCCCATCTA 2340
DB 2970 ACCAACTTAGAAGCATATTAAAGCAGAGACGTTACTTTGCTGACTAAGTTCTGTCTA 3029
QY 2341 GTCAGGCTATGTTTTTCCAGTGGTCAATGATGATGAGAGTGGAGCTGGAAGAA 2400
DB 3030 GTCAAACCTATGGTTTTTCCAGTAGTATATATGATGT--GAGTTGAACATATAAGAAA 3087
QY 2401 GCTGAGCACTGAAGAAATGATGCTTTTGAACCTGTGGTTTGGAGAGACCTTTGAGAGTC 2460
DB 3088 GCTGAGCACCACAAAGATTTGCTTTTGAATTTGGTTTGGAGAGTCTCTTGAAGATC 3147
QY 2461 CTTTGA-CTGCAAGGAGATCCAAACAGTCCATTTCTGAAGAGATCAGCCCTGGGATTC 2519
DB 3148 CTTTGAACCTTGAAGGAGATCCAAACAGTCCATCTTAAAGAAATCAGTCTGAAATATTC 3207
QY 2520 TTTGGAAGGAATGATGCTTAAAGCTGAACCTCCAGTACTTTGGCCACCTGA--TCAGAAG 2577
DB 3208 ATTGGAAGGACTGATGCTGAAATGAAGATTAACGTTTTTGGATCTACCTTAATGCAAGA 3267
QY 2578 GCTGACTCACTGAAAGACCTGTGCTGGAGGATTTGGGGCAGAGGAGGAGGGA 2637
DB 3268 GCCAACTCACTAGAAAAGACC-CATGTTGGCAAAAATTAAGCCAGGAAGAGAGTGA 3326

```

```

QY 2638 CGACAGAGGATGAGATGGCTGGATGGCATCACTGACTCGATGACGCTGAGTCTGGGTGA 2697
DB 3327 TGACAGAGGATGAGATGGTGGATGGCATCGTTGACTGATGATGATGATGATGATGATCA 3386
QY 2698 CTCTGAGTGGTGTATGGACAGGGAGGCTGCTCTCGCGCGATTCATGGGGTCAAAAAG 2757
DB 3387 TTCCGGGAGACAGCAAAAGGACAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3446
QY 2758 AGTTGGACACGACTGAGCACTGAAC 2783
DB 3447 AGTCGGTCTCAAATGATGACTAAAC 3472

RESULT 6
US-09-876-143-1493/c
; Sequence 1493, Application US/09876143
; Publication No. US20040081958A1
; GENERAL INFORMATION:
; APPLICANT: Infigen Inc.
; APPLICANT: EILERTSEN, KENNETH J.
; APPLICANT: PEISTER-GENSKOW, MARTHA
; APPLICANT: CHILDS, LYNETTE
; APPLICANT: FORSYTHE, TODD
; APPLICANT: BISHOP, MICHAEL D.
; TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING
; FILE REFERENCE: 028040-0202
; CURRENT APPLICATION NUMBER: US/09/876,143
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/209,874
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 1744
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 1493
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Bovine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(908)
; OTHER INFORMATION: n is a, c, g, or t
US-09-876-143-1493

```

```

Query Match      11.6%; Score 413.4; DB 1; Length 908;
Best Local Similarity 85.6%; Pred. No. 18;
Matches 501; Conservative 0; Mismatches 68; Indels 16; Gaps 4;

QY 1663 CATGAAAAAGCAAGAGAGTCCAGAAAAACATCTATTCTGCTTTATTGACTATGCAAA 1722
DB 653 CATTGAAAAAGCAAGAGAGTCCAGAAAAACATCTATTCTGCTTTATTGATTGCAAA 594
QY 1723 AGCCCTT--GACTGTGGGGTCAATAACTGTGAAAAATCTGAAAGGATGGAATA 1780
DB 593 AGGCTTTTGACTTGTGTGGATCAATAAATCTTATGAAAAATCTGAAAGAGATGGGATA 534
QY 1781 CCAGACCACTGACCTGACCTCTTGAATAATTTGATGAGGTGAGGAGCAAGTAGA 1840
DB 533 CCAGACCACTGATGCTCTTGAATAACCTATATGAGTCAAGAGCAAGTAGA 474
QY 1841 ACTGACATGGAACAAACAGACTGGTTCCAGTAAAAAGGAGTATGTCAGGCTGTATA 1900
DB 473 ACTGACATGGAACAAACAGACTGGGTCCAAATTTGAAAAAGGAGTACATCAAGCTGTCTA 414
QY 1901 TTGTCACCCGCTGTTTAACTTCTATGCAAG-ACATCATGAGAAAGCTGGCTGGA 1959
DB 413 TTGTCACCCGCTGTTTAACTTCTATGCAAGTACATCATGAGAAATGCTGGGCTGGAG 354
QY 1960 GAAGCAAGCTGGAATCAAGATTGCCGGGAGAAATAGCAATAACCTCAGATATGCAGAT 2019
DB 353 AAAGCAAGCTGATTCAGATTGCCGGGAGAAATATCAATAACCTCAGATATGCAGAT 294
QY 2020 GATACCACCTTTATGGCAGAAAGTGAAGAGAACTAAAAAGCCCTCTTGTATGAGGCTGAA 2079

```



```

; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; PRIORITY FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7158
; LENGTH: 436
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 31-LIB2809-013-Q1-E1-H3
US-09-960-352-7158

Query Match          9.7%; Score 347.4; DB 1; Length 436;
Best Local Similarity 92.8%; Pred. No. 54;
Matches 385; Conservative 0; Mismatches 26; Indels 4; Gaps 2;

QY 1931 GAGACATCATGAGAAAGCGCTGGGCTGGAAGAGACAAAGCTGGAATCAAGATTGCCGGGA 1990
      |||
Db 24 GATACATCATGAAAACCCCTGGGCTGGAGGAAGCTAAAAGCTGGAATCAAGATTGCCAGGA 83
      |||

QY 1991 GAAATGCAATTAACCTCAGATATGCGAGATGATACACCCCTTATGGCAGAAAGTGAAGAGG 2050
      |||
Db 84 GAAATATCAATAAGGCTTGATATGCGAGATGACACCACCCTTATGGCAGAAAGTGAAGAG 143
      |||

QY 2051 AACTTAAAGACCCCTCTGTATGTAAGGTGAAAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCA 2110
      |||
Db 144 AACTTAAAGACCCCTCTGTATGTAAGGTGTAAGAGAGAGTGAAGAAAGTTGGCTTAAAGCTCA 203
      |||

QY 2111 ACATTGAGAAAAAGAGATCATGGCATCTGGTCCCATCACTTCATCGGAAATAGATGGGG 2170
      |||
Db 204 ACATTTAGAAAACTAAGATCATGGCATCTGGTCCCATCACTTCATCGGAAATAGATGGGG 263
      |||

QY 2171 AAACAGTGGAAACAGTGTGAGAGCTTTATTTTGGGGGGCTCCAAATCACTGCAGATGG 2230
      |||
Db 264 AAACAGTGGAAACAGTGTGAGAGCTTTA--TTTGGGGGGGCTCCAAATCACTGCAGATGG 321
      |||

QY 2231 TGACTGCAGCCATGAAATTAAGAGACACTTACTCTTTGGAAGAAAGTTA--ACCAACCT 2288
      |||
Db 322 TGATTCAGCCATGAAATTAAGAGCGCTACTCTCTTGGAGGAAGATTATGACCAACCT 381
      |||

QY 2289 AGATAGCATATTGAAAGCAGAGACATTAACCTTGGCCCAACAAAGCCCCCATCTAGTC 2343
      |||
Db 382 AGATAGCATATTGAAAGCAGAGACATTAACCTTGGCCCAACAAAGGTCCTAGTC 436
      |||

RESULT 15
US-10-109-551-3
; Sequence 3, Application US/10109551
; Publication No. US20020194635A1
; GENERAL INFORMATION:
; APPLICANT: DUNNE, PATRICK W.
; APPLICANT: PEDRAHITA, JORGE
; TITLE OF INVENTION: TRANSGENIC ANIMALS RESISTANT TO TRANSMISSIBLE
; TITLE OF INVENTION: SPONGIFORM ENCEPHALOPATHIES
; FILE REFERENCE: TAMK:207US
; CURRENT APPLICATION NUMBER: US/10/109,551
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/280,549
; PRIOR FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 31412
; TYPE: DNA
; ORGANISM: Ovis aries
US-10-109-551-3

```

```
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (383)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 35-LIB3058-006-Q1-K1-A12
US-09-960-352-8154

Query Match          9.4%; Score 335.8; DB 1; Length 428;
Best Local Similarity 86.4%; Pred. No. 62;
Matches 370; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 2308 AGAGACATTACTTGGCCAAACAAAGCCCATCTAGTCAAGGCTATGGTTTTTCCAGTGGTC 2367
Db 1 AGAGACATTACCTTGGCCAAACAAAGGTCGGICTAGTAAGGCTATGTTTTTCCAGTGGTC 60
QY 2368 ATGTATGATGTGAGAGTTGGACTGTGAAGAACTGAGCACTGAGAAATGTGCTTTT 2427
Db 61 ATGTATGATGTGAGAAATTTGGACTATAAAGAATCTGAGCACCAAGAATAATGTGCTTTT 120
QY 2428 GAACTGTGGTCTTGGAGAAGACTCTTGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAG 2487
Db 121 GAACTGTGGTCTTGGAGAAGACTCTTGAGAGTGGCTTGGACTGCAAGGAGATCCAAACAG 180
QY 2488 TCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTTGAAGAAATGATGCTAAAGCTGAAA 2547
Db 181 TCCATTCTGAAGAGATCAGTCTCTGAATATTCATTGGAAGGACTGATGTTGNAAGCTGAAG 240
QY 2548 CTCAGTACTTTGGCCACCTGATCAGAGAGCTGACTCACTTGGAAAGACCCCTGATGCTG 2607
Db 241 CTCAGTACTTTGGCCACCTGATGTGAAGACTGACTCACTTTGAAAGCCCATGATCCTG 300
QY 2608 GGAGGGATTTGGGGCAGGAGAGAGAGGGGCGACAGAGGATGAGATGGCTGGATGGCATC 2667
Db 301 GGAAAGATTAAAGTGAATGAGGAGAGAAAGGATGACAGAAATGAGATGGTTGGATGGCATC 360
QY 2668 ACTGACTCGATGGAGCTGAGTCTGGTGAATCTCTGAGTTGTTGATGGACAGGAGGCC 2727
Db 361 ACCGACTCAATGGACATGGGTTNGGTCGACTCCAGAAAGTTGGTGCATGGATAGCGAAGCC 420
QY 2728 TGTCTGTC 2735
Db 421 TGGGCTGC 428

RESULT 17
US-09-960-352-1308/c
; Sequence 1308, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Ningbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 1308
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (384)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 06-LIB34-016-Q1-E1-B9
US-09-960-352-1308

Query Match          9.3%; Score 333; DB 1; Length 406;
Best Local Similarity 91.9%; Pred. No. 66;
Matches 351; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
```

```
QY 2405 AGACTGAAGAATTGATGCTTTTGAACCTGTGTGTGGAGAAGACTCTTGAGAGTCCCTT 2464
Db 405 AGTCCCGAAGAATTGATGCTTTTNGAATCTGTGTGTGGAGAAGACTCTTTGAGTGTCCCTT 347
QY 2465 GGACTGCAAGGAGATCCAAACCAAGTCCATTCTCTGAAGAGATCAGCCCTGGGATTTCTTTGG 2524
Db 346 GGACTGCAAGGAGATCCAAACCAAGTCCATTCTCTGAAGAGATCAGCCCTGGGATTTCTTTGG 287
QY 2525 AGGAATGATGTAAGCTGAAACCTCCAGTACTTTGGCCACCTGATCAGAAAGAGTGAAT 2584
Db 286 AGGAATGATGCTGAAGATGAAACCTCCAGTACTTTGGCCACCTGATCAGAAAGAGTGGCT 227
QY 2585 CACTGAAAAGACCCTGATGCTGGGAGGATTTGGGGCAGGAGAGAAAGGGACACACAGA 2644
Db 226 CATTGAAAAGACTCTGATGCTGGAGGATTTGGGGCAGGAGAGAAAGGGACGCCAGA 167
QY 2645 GGATGAGATGGTGGATGGCATCACTGACTCGATGGAAGTGTGAGTGTGGAATCTCTGG 2704
Db 166 GGATGAGATGGTGGATGGCATCGTCTGACTCGATGGAAGTGTGAGTGTGGAATCTCTGG 107
QY 2705 AGTTGGTGTGACAGGGAGGCTGTCTCTGGCGGATTCATGGGTGCACAAAGAGTTGA 2764
Db 106 AGTTGGTGTGACAGGGAGGCTGTCTCTGGCGGATTCATGGGTGCACAAAGAGTTGA 2764
QY 2765 CACGACTGAGCAACTGAACCTGA 2786
Db 46 TAGGGCTGAGCGACTGAACCTGA 25

RESULT 18
US-09-960-352-8311/c
; Sequence 8311, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Ningbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 8311
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 36-LIB188-014-Q1-E1-A12
US-09-960-352-8311

Query Match          9.1%; Score 324; DB 1; Length 423;
Best Local Similarity 86.7%; Pred. No. 70;
Matches 357; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 2380 GAGAGTTGAGCTGTGAAGAAAGCTGAGCACTGAGCAAGTAATGATGCTTTTGAAGTGTGT 2439
Db 423 GAGAGTTGAGCTGTGAAGAAAGCTGAGCGCCCAAGAAATGATGCTTTTGAAGTGTGT 364
QY 2440 TGGAGAAGACTCTTGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCTGAAG 2499
Db 363 TGGAGAAGACTCTTGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCTGAAG 304
QY 2500 GAGATCAGCCCTGGGATTTCTTTTGAAGGAATGATGCTAAAGCTGAAATCTCCAGTACTTT 2559
Db 303 GAGATCAGCCCTGGGATTTCTTTTGAAGGAATGATGCTAAAGCTGAAATCTCCAACTTT 244
QY 2560 GGCCACTGATCAGAAAGACTGACTGGAAGAAAGACCCCTGATGCTGGGAGGATTTGG 2619
Db 243 GGTCACTGATGGGAAAGAACTGACTGATCTGAAAAGACCCCTGATGCTGGGAAAGATTGA 184
QY 2620 GGAGGAGGAGAGGGGACGACAGAGGATGAGATGGCTGGATGGCATCACTGATCGATG 2679
```


NAME/KEY: unsure
LOCATION: (31)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: 61-LIB3058-039-Q1-K1-H2
US-09-960-352-14308

Query Match 9.0%; Score 320.4; DB 1; Length 409;

Best Local Similarity 89.3%; Pred. No. 74;
Matches 366; Conservative 0; Mismatches 42; Indels 2; Gaps 2;

Qy 1748 AAACGTGGGAAATCTCTGAAGGGATGGGAATACACAGACACCTGACCTGACCTTTGAAA 1807
Db AAACGTGGGAAATCTCTGAAGGGATGGGAATACACAGACACCTGACCTGACCTTTGAAA 350
Qy 1808 AATTTGTATCAGGTGAGGAAGCAACAGTTAGAACTGGAATGGAACACAGACTGGTTC 1867
Db AATTTGTATCAGGTGAGGAAGCAACAGTTAGAACTGGAATGGAACACAGACTGGTTC 290
Qy 1868 CAAGTAGGAAGAGATCTCAAGGCTGTATTTGTCACCGGCTTGTAACTTCTAT 1927
Db CAAGTAGGAAGAGATCTCAAGGCTGTATTTGTCACCGGCTTGTAACTTCTAT 230
Qy 1928 GCAGAG-ACATCATGAAACGCTGGGCTGGAAGAACACAACTGGAATCAAGATTGCC 1986
Db GCAGAGTCCATGATGAAACCTCTGGGCTGGAAGAACACAACTGGAATCAAGATTGCC 170
Qy 1987 GGGAGAAATAGCAATACCTCAGATATGAGATGATACACCTTATGGCAGAAAGTGA 2046
Db GGGAGAAATAGCAATACCTCAGATATGAGATGATACACCTTATGGCAGAAAGTGA 110
Qy 2047 GAGAACTTAAAGACCTCTTGATGAGGTGGAAGAGAGAGTGAAGAAAGTTGGCTTAAAG 2106
Db GAGAACTTAAAGACCTCTTGATGAGGTGGAAGAGAGAGTGAAGAAAGTTGGCTTAAAG 50
Qy 2107 CTCAACTTCAGAAACGAGATCATGGCTCTGCTGCCATCACTTCATG 2156
Db CTTCACCTTCAGAAATG-NGATCATGCATCGCTCCCTCCGTTCCATG 1

RESULT 22

US-09-960-352-9545
Sequence 9545, Application US/09960352
Patent No. US2002037139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)C

CURRENT APPLICATION NUMBER: US/09/960.352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 9545

LENGTH: 425

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 41-LIB3058-043-Q1-K1-C10

US-09-960-352-9545

Query Match 8.8%; Score 313.4; DB 1; Length 425;

Best Local Similarity 87.4%; Pred. No. 77;
Matches 375; Conservative 0; Mismatches 46; Indels 8; Gaps 3;

Qy 2073 GGTGAAGAGGAGGTGAAAAGTTGGCTTAAAGCTCAACATTGAGAAACGAAAGATCAT 2132
Db GGTGAAGAGGAGGTGAAAAGTTGGCTTAAAGCTCAACATTGAGAAACGAAAGATCAT 62
Qy 2133 GCATCTGTCCTCCATCTCATGGAATATAGATGGGAAACAGTGGAAACAGTGTGAGA 2192
Db GCACGTGATCCCATTTATTCATGCAATATAGATGGGATCAATGGAACAGTGTGAGA 122

Qy 2193 CTTTATTTTTTGGGGGCTCCAAATCACTGAGTGGTACTGACGACCATGAATTTAAA 2252
Db CTTTATGTTT--TGGGCTCCAAATCACTGAGTGGTGAATGAGCCATGAATTTAAA 180
Qy 2253 AGACACTTACTCCTTTGGAAGAAAAGTTA--ACCAACCTTAGATAGCATATTGAAAAGCAGA 2310
Db TGACTCTTCTCTTTGGAAGAAAAGCTATGACCGAGCTAGACAGCATATTAAAAAGCAGA 240
Qy 2311 GACATTACTTGGCCACAAAGCCCACTAGTCAAGGCTATGGTTTTCAGTGGTTCATG 2370
Db AACATTACTTTTACCACAAAGGTCATCTATTTCAAAGCTA----TTTTCTAGTAGTCATG 296
Qy 2371 TATGATGTGAGAGTTGACTGTGAAGAAAGCTGAGCACTGAAGAAATGATGCTTTTGA 2430
Db TATGTTGTGAGAGTTGACTATAAAGAAAGCTGAGTCTCAAGAACTGATGTTTGA 356
Qy 2431 CTGTGGTGTGGAAGAACTCTTGAAGTCCCTTGGACTGCAAGAGAGATCAACAGTCC 2490
Db CTGCGGTGTGTGGAAGAACTCTGGAAGTCCCTTGAAGTCCCTTGAAGAGATCAACAGTCA 416
Qy 2491 ATTCTGAAG 2499
Db ATCTTAAG 425

RESULT 23

US-09-960-352-3391

Sequence 3391, Application US/09960352

Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengbing

APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21(10298)C

CURRENT APPLICATION NUMBER: US/09/960.352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 3391

LENGTH: 422

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 15-LIB3058-036-Q1-K1-D11

US-09-960-352-3391

Query Match 8.6%; Score 307.8; DB 1; Length 422;

Best Local Similarity 86.0%; Pred. No. 81;
Matches 363; Conservative 0; Mismatches 57; Indels 2; Gaps 2;

Qy 1714 CTATGCAAAAGCCTTTGACTGTGGGGTCAATAAATGAGTGGGAAATTTCTGAAAGGAT 1773
Db CTGCGTGAATAACCTTTGACTGTGTAGATCAACAATCTGTGGAATAATTTGTTAAAGAGAT 60
Qy 1774 GGGATACCAAGACCACTGACTGCTTGAATAATTTGATGAGGTGAGGAGCAAC 1833
Db GGGATACCAAGACCACTGACTGCTTGAATAATTTGATGAGGTGAGGAGCAAC 120
Qy 1834 AGTTAGAACTGGACATGGAAACAAACAGACTGGTTCAGGTAGGAAAGAGTATGTCAAG 1893
Db AGTTAGAACTGGACATGGAAACAAACAGACTGGTTCAGGTAGGAGTATGTCAAG 180
Qy 1894 CTGTATTTGTACCCGGCTTGTTTAACTTCTATGAGAG-ACATCATGAGAACTCG 1952
Db CAATATTTGTCACTCCCTGCTTATTTAACTTATATGAGAGTACATCATGAGAAATCTG 240
Qy 1953 GCTGGAAGAGCAACAGCTGGAATCAAGATTGC-CCGGAGAAATAGCAATTAACCTCAGAT 2011
Db GCTGGAATGAACAAAGCTGGAATCAAGATTGCTGGGAGAAATATCAATTAACCTCAGAT 300
Qy 2012 ATGCAGATGATACCACTTATGGCAGAAAGTGAAGAGGAACTTAAAGACCTCTTGATGA 2071

Query Match 8.3%; Score 296; DB 1; Length 410;
Best Local Similarity 86.9%; Pred. No. 92;
Matches 358; Conservative 0; Mismatches 50; Indels 4; Gaps 3;

QY 593 AATGGAGTGGTCAACAAGAGATGCGAAGAGTGAATGCAACATTTAGGAATCAGCG 652
DB 410 AATGGAGGTTCAACAGGAGATGCGAAGATGAACATCAACATTTAGGGATCACTG 351
QY 653 AACTAAAATGAGTGGGAATGGGTGAATTTAACTCAGATGACCAT-ATATCTACTACTGC 711
DB 350 AAATAAATGAAGTGGTGAATGGGTGAATATAACGAGATGACCATTAATATCTACTACTGT 291
QY 712 GGGCAGAAATCCCTCAAGAAATGGAGTAGCCATCATGCTCAACAAAAGAGTCCGAAT 771
DB 290 GGGCAAGAAATCCCTTGAAGAAATGGAGTAGCCATCATATCAACGTAAAGTCTGACAT 231
QY 772 GCAATCTGGATGATCTCAAAAACGACAGAAATGATCTCTGTTTTCAGAGGCAAA 831
DB 230 GCA-ACCTGGATGCAATCTCAAAAACGACAGAAATGATTTCT-ATTCAATCCAAAGGCAAA 174
QY 832 CCATTCATATCAGTAATCCAGTCTATGCCCAACCAAGTAAATGCTGAAGAGCTGAA 891
DB 173 CCATTCATATCAGTAATCCAGTCTATGCCCAACCAAGTAAATGCTGAAGAGCTGAT 114
QY 892 GTTGAAGGCTCTATGAAGACCTCAAGAGCTTTTGAAGTAAACCCCAAAAAGATGTC 951
DB 113 GGTGAATGGTCTATGAAGACCTCAAGAGCTTTTGAAGTAAACCCCAAAAATATGTC 54
QY 952 CTTCTCATATAGGGGAGTGAATGCAAAAGTAGGAGCAAGAACACCTG 1003
DB 53 CTTTTCATATAGTGGAGTGGATGCAAAAGTAGGAATACAGAGATATATG 2

RESULT 27
US-09-960-352-3913/c
; Sequence 3913, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3913
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 17-LIB34-010-Q1-E2-E1
US-09-960-352-3913

Query Match 8.1%; Score 289.4; DB 1; Length 356;
Best Local Similarity 91.6%; Pred. No. 1.le-02;
Matches 327; Conservative 0; Mismatches 26; Indels 4; Gaps 2;

QY 2183 CAGTGTGACACTTTATTTTGGGGGGCTCCAAATCACTGCAGATGGTGAAGTGAAGCA 2242
DB 356 CAGTGTGACACTTTA--TTTGGGGGGCTTCAATCACTGCAGATGGTGAAGCA 299
QY 2243 TCAAAATTAAGACACTTACTCTCTTGGGAAGAAAGT--TAACCAACCTAGATAGCATATT 2300
DB 298 TCAAAATTAAGACACTTACTCTCTTGGGAAGAAAGT--TAACCAACCTAGATAGCATATT 239
QY 2301 GAAAGCAGAGACATTAATCTTCCCAAGACCCCTAGTCAAGGCTATGTTTTTCC 2360
DB 238 CAAAGCAGAAACATTAATCTTCCCAAGACCCCTAGTCTTAATCAAGGCTATGTTTTTCC 179
QY 2361 AGTGGTGCATGTATGTAGTGTGAGATGGATGTGGAAGAAAGCTGACACTGAAGAATTGA 2420

DB 178 AGTGGTGCATGTATGTAGTGTGAAGTGTGAAGAAAGCTGAGTCTGAAGAATTGA 119
QY 2421 TCGTTTGAACATGTTGGTGTGGAGAGACTCTTGAGAGTCCCTTGAGTCAAGGAGATC 2480
DB 118 TCGTTTGAACATGTTGGTGTGGAGAGACTCTTGAGAGTCCCTTGAGTCAAGGAGATC 59
QY 2481 CAACCAAGTCCATTCGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCT 2537
DB 58 CAACCAAGTCCATTCGAAGGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATGCT 2

RESULT 28
US-09-960-352-13876
; Sequence 13876, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13876
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 59-LIB34-027-Q1-E1-E4
US-09-960-352-13876

Query Match 8.0%; Score 286; DB 1; Length 413;
Best Local Similarity 87.6%; Pred. No. 1e+02;
Matches 345; Conservative 0; Mismatches 45; Indels 4; Gaps 3;

QY 1889 CAAGGCTGTATTTGTACCCGGCTTTTAACTTCTTATGACAG-ACATCATGAGAAC 1947
DB 1 CAAGGCTGTATTTGTACCCGGCTTTTAACTTCTTATGACAGTACATCATGCGAAT 60
QY 1948 GCTGGGCTGGAAGAACACAAAGCTGAATCAAGATTCGCGGAGAAATAGCAATACCTC 2007
DB 61 CTGGGCTGTGATGAGGACCAACCTGGAATCAAGATTCGCGGAGAAATATCAATACCTC 120
QY 2008 AGATATGACAGATGATACCCCTTATGCGAGAAAGTGAAGAGAACTAAAAGCCTCTTG 2067
DB 121 AGATATGACAGATGACACCATCTCTTAATGGCAAAAGTGTAGAGAACTAAAGAGCCTCTG 180
QY 2068 ATGAAGGTGAAGAGGAGAGTGAAGAAAGTGGCTTAAAGCTCAACATTCAGAAAAAGCAAG 2127
DB 181 ATGAAGGTGAAGAGGAGAGTGAAGAAAGTGGCTTAAAGTTCACATTCAGAAAAAGCAAG 240
QY 2128 ATCATGGCATCTGGT-CCCATCACTTCATGGGAATAGATGGGAAACAGTGGAAACAGT 2186
DB 241 ATCATAGCATCTGGTCCCATCATTTTCATGGCAAAATATATGGAGAAATATATGGAAACAG 300
QY 2187 GTCAGACTTTATTTTGGGGGGCTCCAAATCACTGCAGATGGTGCATGCGCCATGAA 2246
DB 301 GACAGATTTTATTTCT--TGGGCTCCAAATCACTGTGATGGTGTGACTGCACCCATGA 358
QY 2247 ATTAAGAGACACTTACTCTCTTGGGAAGAAAGTTA 2280
DB 359 ATTAAGAGACACTTATCTCTTGGGAAGAAAGTTA 392

RESULT 29
US-09-960-352-9592
; Sequence 9592, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 9592
LENGTH: 405
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 41-LIB34-030-Q1-E1-C2
US-09-960-352-9592

Query Match
Best Local Similarity 8.0%; Score 285.2; DB 1; Length 405;
Matches 334; Conservative 0; Mismatches 56; Indels 8; Gaps 1;

QY 2371 TATGGATCTGAGAGTTGAGCTCTGAAGAAAGCTGAGCACTGAAGAAATTGATGCTTTTGA 2430
Db 14 TGTTCATTTTGTGAGACTATAAAGAAAGCTGAGCACTGAAGAAATTGATGCTTTTGA 73
QY 2431 CTGTGGTCTGGAGAGACTCTTGAGAGTCCCTTGGAGCTGCAAGAGATCCCAACAGTCC 2490
Db 74 CTGTGGTCTGGAGAGACTCTTAAGAGTCCCTTGGAGCTGCAAGAGATCCCAACAGTCC 133
QY 2491 ATTCCTGAAGAGATCAGCCCTGGGATTTCTTGAAGAAATGATGCTAAAGCTGAACCTC 2550
Db 134 ATCTTAAGAGAGATCAGCTCTAGTGTTCATTTGAAGAGCTGAAGTTGAAGCTGAACCTC 193
QY 2551 CAGTACTTGGCCCTGATCAGAGAGCTGACTCACTGGAAGAAAGCCCTGATGCTGGGA 2610
Db 194 CAATACTTGGCCCTGATCAGAGAGCTGACTCACTTGAAGAAAGCCCTGGAAGCTGGGA 253
QY 2611 GGGATTTGGGCGAGAGAGAGAGGAGCAGAGAGTGAAGTGGCTGGAGTGGATGCTCACT 2670
Db 254 AAGATTGGGCGAGAGAGAGAGGAGCAACAGAGGCGAGATGTTGGATAGCATCACC 313
QY 2671 GACTCGATGAGCTGAGTCTGGGTGAATCTCTGGAGTTGGTGAAGACAGGGAGGCTGT 2730
Db 314 GACTCAATGAGCATAGTTTGGGTGAGTCTGGGAGTTGGTGAAGACAGGGAGCTT-- 371
QY 2731 CCTGGCGATTCATGGGTCACAAAGAGTTGACAGAC 2770
Db 372 -----GGGGTTTCATGGGGTCTCAGAGAGTCAGACGAC 405

RESULT 30
US-09-960-352-13585/c
Sequence 13585, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 13585
LENGTH: 407
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 58-LIB3058-010-Q1-K1-G6
US-09-960-352-13585

Query Match
Best Local Similarity 8.0%; Score 284.5; DB 1; Length 407;
Matches 334; Conservative 0; Mismatches 56; Indels 8; Gaps 1;

Matches 337; Conservative 0; Mismatches 70; Indels 1; Gaps 1;
QY 2377 TGTGAGAGTTGAGCTCTGAAGAAAGCTGAGCACTGAAGAAATTGATGCTTTTGAACCTGG 2436
Db 407 TGTGAGAGTTGAGCTCTGAAGAAAGCTGAGCACTGAAGAAATTGATGCTTTTGAACCTGG 349
QY 2437 TGTGAGAGAGACTCTTGAAGAGTCCCTTGGAGTCTGCAAGAGATCCAAACAGTCCATTCTG 2496
Db 348 TGTGAGAGAGACTCTTGAAGAGTCCCTTGGAGTCTGCAAGAGATCCAAACAGTCCAACTA 289
QY 2497 AAGGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATCTAAAGTGAACCTCCAGTAC 2556
Db 288 AATGAAATCAGTCTAAATATTTCAITTTGGAAGAGTCTGCTGAAGCTGAACCTCCAGTAC 229
QY 2557 TTTGGCCACTGATCAGAGAGTCTGACTCAGTGGAAAGACCTGATCTGCTGGAGGGATT 2616
Db 228 TTTGGCCACTGATGGAAGAACTGACTCTTTTGAAGAACTGATCTGCTGGAGAAAGATT 169
QY 2617 GGGGGCAGAGAGAGAGGGGACGACAGAGGATGAGATGCTGGATGCTGATCACTGACTCG 2676
Db 168 GAAAGCATTAAGGAGAAAGGGACGACAGAGATGATGCTTGAATGGCATCACCGACTCG 109
QY 2677 ATGGAGTGAAGTCTGGTGAAGTCTCTGAGTGGTGGATGAGCAGAGAGGCTGCTGCTGG 2736
Db 108 ATGGAGTGAAGTCTGGTGAAGTCTCTGAGTGGTGGATGAGCAGAGAGGCTGCTGCTGGTATCT 49
QY 2737 GCGATTTCATGGGTCACAAAGAGTTGGACACGACTGAGCACTGAGCACTGAGTGAAC 2784
Db 48 GCGATTCATGGGTCACAAAGAGTCTTGAAGAGTCTTGAAGAGTCTGAGTGAAC 1

RESULT 31
US-09-960-352-13323
Sequence 13323, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 13323
LENGTH: 423
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 57-LIB3057-008-Q1-K1-G2
US-09-960-352-13323

Query Match
Best Local Similarity 7.9%; Score 282.9; DB 1; Length 423;
Matches 356; Conservative 0; Mismatches 66; Indels 7; Gaps 3;

QY 1637 GACCAATTTGCAACATCTCTGTATCTATGCAATGCAATGCAATGCAATGCAATGCAATG 1696
Db 1 GATCAATTTGCAACATCTCTGTATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 60
QY 1697 TATTTCTCTTTTATTTGACTATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 1756
Db 61 TACATCTACTTCACTGACTTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 118
QY 1757 AAAATTTGAAAGGATGGGAATCCAGACCACTGACTGCTTGTGAAATTTGTAT 1816
Db 119 AAAATTTGAAAGGATGGGAATCCAGACCACTGACTGCTTGTGAAATTTGTAT 178
QY 1817 GCAGGTGAGGAGCAACAGATTGAGAACTGGACATGGAACACAGAGTGTTCAGAGTAGGA 1876
Db 179 GCAGGTGAGGAGCAAC-----AGAACTGGACATGGAATTTGAAATTTGCTTCAATTTGGG 234

QY 1877 AAAGGAGTATGTCAGGCTGTATATGTCACCGGCTTCTTTAACTTCTATCGAG-AC 1935
Db 235 AAAGGTTTATGTCAGGCTGCATATGTCACCGCTCTTATTTAACTTATGAGAGTAC 294
QY 1936 ATCATGAGAAACGCTGGGCTGGAAGACACAGCTGGAATCAAGATTGCCGGAGAAAT 1995
Db 295 ATCATGTGAATGCTGAGCTGGATGAATCACAACTGAAATTAGGATTGCAAGGAGAAAT 354
QY 1996 AGCAATACTCAGATGATGAGATGATACACCGCTTATGCGCAGAAAGTGAAGGAACTA 2055
Db 355 ATCGACAGTCTCAGATGATGAGACGATACCAATTTAATGCGAGAAAGTGAAGAGGAACTA 414
QY 2056 AAAAGCCTC 2064
Db 415 AAGATCCTC 423
RESULT 32
US-09-960-352-13570
; Sequence 13570, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13570
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 58-LIB3057-018-Q1-K1-G6
US-09-960-352-13570
Query Match 7.8%; Score 279.1; DB 1; Length 365;
Best Local Similarity 94.2%; Pred. No. 1.2e+02;
Matches 341; Conservative 0; Mismatches 14; Indels 7; Gaps 5;
QY 2024 CCACCTTATGGCAGAAAGTGAAGGAACTTAAAGCCCTTGTATGAAGGTGAAGAGG 2083
Db 5 CCACCTTGTGGCAG-AAGTGAAGGAACTTAAAGCCCT-TTGATGAAAGTGAAGA-G 61
QY 2084 AAGTGAAAAGTTGGCTTAAAGCTCAACATTGAGAAAACGAGATCATGGCATCTGGTC 2143
Db 62 AAGTGAAAAGTTGGCTTAAAGCTCAACATTGAGAAAACGAGATCATGGCATCTGGTC 121
QY 2144 CCATCATTTCATGGGAAATAGATGGGAAACAGTGGAAACAGTGTCTAGACTTTATTTT 2203
Db 122 CCATCATTTCATGGGAAATAGATGGGAAACAGTGGAAACAGTGTCTAGACTTTATTTT 181
QY 2204 GGGGGCTCCAAATCATTGAGATGTTGATGCTGAGCCATGAATTAAGACACTTACT 2263
Db 182 --TGGGCTCCAAATCATTGAGATGTTGATGCTGAGCCATGAATTAAGACACTTACT 239
QY 2264 CTTTGGAGAAAAGTTA--ACCACTAGATAGATATTTGAAAACAGACACATTAACCTT 2321
Db 240 CTTTGGAGAAAAGTTAGACCCNACCTAGATAGATATTTGAAAACAGACACATTAACCTT 299
QY 2322 GCCAACAAAGCCCACTAGTCAAGGCTATGGTTTTCCAGTGTGATGATGATGA 2381
Db 300 GCCAACAAAGGTCCTCTAGTCAAGGCTATGGTTTTCCAGTGTGATGATGATGA 359
QY 2382 GA 2383
Db 360 GA 361
RESULT 33

US-09-960-352-10238
; Sequence 10238, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 10238
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 44-LIB3058-004-Q1-K1-C8
US-09-960-352-10238
Query Match 7.7%; Score 275.7; DB 1; Length 448;
Best Local Similarity 88.7%; Pred. No. 1.1e+02;
Matches 330; Conservative 0; Mismatches 33; Indels 9; Gaps 3;
QY 1831 AACAGTTAGAACTGGAACATGGAACACAGACTGGTTCCAAGTAGGAAAAGGATGTC 1890
Db 78 AACAGTTAGAACTGGAACATGGAACACAGACTGGTTCCAAGTAGGAAAAGGATGTC 137
QY 1891 AGCGTGTATATGTCACCGGCTTCTTTAACTTCTATGAGAG-ACATCATGAGAAAAGC 1949
Db 138 ATGTTGTATGTCACCGCTTCTTTAACTTCTATGAGAGTACATACAAAATAC 197
QY 1950 TGGGCTGGAAGAACACAGCTGGAATCAAGATTCGGGAGAAATAGCAATACTCTAG 2009
Db 198 TGGGCTGGAATGAAGCACAGCTGGAATCAAGATTCAGGACAAAATATCAATACTCTAG 257
QY 2010 ATATGAGATGATACCACTTATGCGAAGAGTGAAGAGGAACTAAAAGCCTCTTAT 2069
Db 258 ATAGCAGATGATACCACTTATGCGAAGAGTGAAGAGGAACT-AGAGCCTCTTAT 316
QY 2070 GAAAGTGAAGAGGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 317
Db 317 GAAAGTGAAGAGGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGA 376
QY 2130 CATGGCATCTGTCCTCATCATCTTCTATGGGAAATAG-----ATGGGNAACAGTGGAA 2182
Db 377 CATGGCATCTGTCCTCATCATCTTCTATGGGAAATAG-----ATGGGNAACAGTGGAA 436
QY 2183 CATGGTCAGACT 2194
Db 437 CAATGGCAGACT 448
RESULT 34
US-09-960-352-6765
; Sequence 6765, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6765
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Bos taurus

Db 37 TTAGATCTGATCAACAGAGTGCTGA 12

RESULT 37

US-09-960-352-7912

Sequence 7912, Application US/09960352

Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengbing

APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND FAT DEPOSITION

FILE REFERENCE: 16511.006/37-21 (10298)C

CURRENT APPLICATION NUMBER: US/09/960,352

CURRENT FILING DATE: 2001-09-24

NUMBER OF SEQ ID NOS: 15112

SEQ ID NO 7912

LENGTH: 449

TYPE: DNA

ORGANISM: Bos taurus

OTHER INFORMATION: Clone ID: 34-LIB3058-004-Q1-K1-A6

US-09-960-352-7912

Query Match 7.6%; Score 270.2; DB 1; Length 449;

Best Local Similarity 82.1%; Pred. No. 1.1e+02; Indels 22; Gaps 5;

Matches 366; Conservative 0; Mismatches 58;

QY 1666 GGAAGGAGAGAGAGTCCAGAAAAACATCTATTCTGTTTATTGACTATGCAAAAGC 1725

DB 12 GGAAGGAGAGAGAGATTTAAAAAATACTACTTCTGTGTCATTGACTACACTAAAGC 71

QY 1726 CTTTGACTGTGGGGTCACAATAAACHGTGGAAAAATCTGAAGGGATGGGATACCAGA 1785

DB 72 TTTTATAT---GGATCACAAACACTGTGAAAAATCTTAAAGAGGTGGGATATCGA 127

QY 1786 CCACCTCACCTGACTCTTTGAAAAATTTGATGCAGGTGAGAAAGCAACAGTTAGAACTGG 1845

DB 128 CCACCTTACCTGACAC---AGAAATCTGTGTGAGGACAGAAACACAGTTAGTACTGG 184

QY 1846 ACATGGAAACACAGACTGGTTCCAGTAGAGAAAGAG-----TATGTCAAG 1892

DB 185 ACATGGAAACACAGAGTGGTTCAAAATGGGAAAGAGTCAAGGACAAATATATGTCAAG 244

QY 1893 GCTGTATTGTCCACCGGCTGTTTAACTTCTATGCAGAG-ACATCATGAGAAAGCTG 1951

DB 245 GCTATATTGTCACTCCCTGCTTTTAACTTATGCAAGTATGCAAGTACATGAGAAATGCCA 304

QY 1952 GGCTGGAGAGACCAAGCTGGATCAAGATTGCCGGAGAAAATAGCAATAACCTCAGAT 2011

DB 305 GACTGGATGAAGCAAGCTGGAATCAAGATTGCTGGAGAAAATATCAATAACCTCAGAT 364

QY 2012 ATGCAGATGATACCACTTTATGCGAAGTGAAGGAACTTAAAGGCTCTTCATGA 2071

DB 365 ATGCAGATGATGCCA-CTTTATGTGAAGTGAAGGAACTTAAAGGCTCTTCATGA 423

QY 2072 AGGTGAAAGAGGAGAGTCAAAAGCTT 2097

DB 424 AGGTGAAAGAGGAGAGTCAAAAGCT 449

RESULT 38

US-09-960-352-11744/c

Sequence 11744, Application US/09960352

Patent No. US20020137139A1

GENERAL INFORMATION:

APPLICANT: Warren, Wesley C.

APPLICANT: Tao, Nengbing

APPLICANT: Byatt, John C.

APPLICANT: Mathialagan, Nagappan

TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND

	Best Local Similarity	91.88	Pred. No. 1.2e+02	Length 492
	Matches	303	Conservative	0
			Mismatches	23
			Indels	4
			Gaps	2
QY	2047	GAGGAAC	TAAAAAGCCTCTTGATGAAGGTGAAAGAGGAGAGTGAAAGAGTTGGCTTAAAG	2106
Db	432	GAAGAACT	CAAAAGCCTCTTGATGAAGGTGAAAGAGGAGAGTGAAAGAGTTGGCTTAAAA	373
QY	2107	CTCAACATT	CAGAAAAACGAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAATAGAT	2166
Db	372	CTCAACATT	CAGAAAAACAAAGATCATGGCATCTGGTCCCATCACTTCATGGGAAATAGAT	313
QY	2167	GGGGAAC	CAGTGAACAGTGCAGACTTTATTTTTTGGGGGCTCCAAATCACATGCAG	2226
Db	312	GGGGAAC	CAGTGAACAAATGTCAGACTTTATTTTTT--TGGGCTCCAAATCACATGCAG	255
QY	2227	ATGGTCACT	GCACCCATGAATTTAAAGACACTTACTCCTTGGAGAAAGTTA--ACCA	2284
Db	254	ATGGTAATT	TGCACCCATGAATTTAAAGATGCTTACTCCTTGGAGAAAGTTATGACCA	195
QY	2285	ACCTAGAT	GACATTTGAAAAGCAGAGACATTACCTTGCACAACAAAGCCCATCTAGTCA	2344
Db	194	ACCTAGAT	GACATTTCAAAAGCAGAGCCATTACTTTGCCAACAAAGGTTCTGTAGTCA	135

;; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
;; FILE REFERENCE: 16511.006/37-21(10298)C
;; CURRENT APPLICATION NUMBER: US/09/960,352
;; CURRENT FILING DATE: 2001-09-24
;; NUMBER OF SEQ ID NOS: 15112
;; SEQ ID NO 14979
;; LENGTH: 423
;; TYPE: DNA
;; ORGANISM: Bos taurus
;; OTHER INFORMATION: Clone ID: 64-LIB3058-007-Q1-K1-H8
US-09-960-352-14979

Query Match 7.2%; Score 258.1; DB 1; Length 423;
Best Local Similarity 91.9%; Pred. No. 1.3e+02;
Matches 283; Conservative 0; Mismatches 24; Indels 1; Gaps 1;
QY 1589 GAATCTCCAGTGTTCAGGCTGTTTGTAGAAAAGTCAGAGAACCCAGAGACCAAAATGGC 1648
DB 320 GAATCTCCAGTGTTCAGGCTGTTTGTAGAAAAGTCAGAGAACCCAGAGATCAAAATGGC 261
QY 1649 AAGATCTCTGTATCATGGAAGAGAGAGATTCAGAGAAACATCTATTTCTGCTTT 1708
DB 260 AAGATCTCTGTATCATGGAAGAGAGATTCAGAGAAACATCTATTTCTGCTTT 201
QY 1709 ATTCACTATGCAAAAGCTTTGACTGTGGGGTCACAAT-AAAGTGTGGAATTTCTGAA 1767
DB 200 ACTGACTTTGCCAAGTCTTTGACTGTGTGATCAATTAAGTGTGGAATTTCTGAA 141
QY 1768 AGGATGGGAATACCAAGACCTGACTGACTCTTGAATAATTTGTATCAGGTAGGA 1827
DB 140 AGAGATGGGAATACCAAGACCTGACTGACTCTTGAATAATTTGTATCAGGTAGGA 81
QY 1828 AGCAACAGTGTAGTGTGATGCAAGCAACAGAGCTGTTTCAAGTAGGAAAGAGATG 1887
DB 80 AGCAACAGTGTAGTGTGATGCAAGCAACAGAGCTGTTTCCATATAGGAAAGAGTACG 21
QY 1888 TCAAGGCT 1895
DB 20 TCAAGGCT 13

RESULT 43
US-09-960-352-149/c
; Sequence 149, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 149
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 01-LIB34-016-Q1-E1-A1
US-09-960-352-149

Query Match 7.2%; Score 257.8; DB 1; Length 417;
Best Local Similarity 82.1%; Pred. No. 1.3e+02;
Matches 330; Conservative 0; Mismatches 62; Indels 10; Gaps 3;
QY 1854 CAACAGACTGTTTCAAGTAGGAAAGAGATGTCTCAGGCTGTATATTTGTCACCG-GC 1912
DB 393 CAATGAGCTGTTTCAAAATGGAAAGAGATACCAAGAGGTGTATATTTGTCACCTTAGC 334
QY 1913 TTGTTTAACTTCTATGCAGACATCATGAGAAAGCTGGGTGGAGAGACCAAGCTG 1972

DB 333 TTATTTAACTTGAAT-----ACATCAGGCAAAATGCCGAGTGGATGAACCAAGCTG 280
QY 1973 GAATCAAGATTGCCGGGAGAAATAGCATTAACCTCAGATATGCAGATGATACCAACCTTA 2032
DB 279 GAGTCAAGATTGCCGGGAGACATATCAACAATCTCAGATATGTAGATATATCAACCTTAA 220
QY 2033 TGGCAGAAAGTGAAGAGCAACTAAAAAGCCTCTTGTGAAGAGTGAAGAGGAGAGTGA 2092
DB 219 TGGCAGAAAGCAAGAGAACTAAGAGCTATTTGTGAAGTGAAGAGGAAAGTGA 160
QY 2093 AAGTTGGCTTAAAGCTCAACATTGAGAAAAGAGATCATGCATCTGTCCCATCACTT 2152
DB 159 AAGCTGGCATAAACTCAAAATTTCAAAAAGTGAATCTGTGCAACTGATCTCATCACTT 100
QY 2153 CATGGAAATAGATGGGAAACAGTGGAAACAGTGTGAGACTTTATTTTGGGGGGCTC 2212
DB 99 CATGCAATAGATGGGAGAAATGGAACAGTGGCAATTTATTTCTT---GGGGCTC 43
QY 2213 CAAAATCACTGCAGATGTGACTGCAGCCATGAAATTTAAAG 2254
DB 42 CAAAATCACTGCAGACAGTGAATGAGCCATGAAATTTAAAG 1

RESULT 44
US-09-960-352-7209
; Sequence 7209, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 7209
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 31-LIB3058-021-Q1-K1-H11
US-09-960-352-7209

Query Match 7.1%; Score 254.8; DB 1; Length 456;
Best Local Similarity 83.6%; Pred. No. 1.3e+02;
Matches 300; Conservative 0; Mismatches 57; Indels 2; Gaps 1;
QY 2431 CTGTGTTTGGAGAGACTCTTGAGAGTCCCTTGAGCTGCAAGGAGATCCCAACAGTCC 2490
DB 1 CTGTGTTTGGAGAGACTCTTGAGAGTCCCTTGAGAGTCCCTTGAGAGATCCCAACAGTCC 60
QY 2491 ATTCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAGGAATGATCTAAAGTGAACCTC 2550
DB 61 ATCTTAAAGAAATCAATCTGTAATTTCACTGGAAGGACTGATGCTGAAGCTGAACTC 120
QY 2551 CAGTATTGGCCACCTGATCAGAGAGTCACTCTGGAAGAGACCTGATGCTGGGA 2610
DB 121 CAGTACTTTGGCCACCTGATGCGAAGAACTGACTCTTTTGAAGAGACCTGATGCTGGGG 180
QY 2611 GGGATTGGGGCAGGAGGAGAGGGGACGACAGAGGATGAGATGGCTGGATCACT 2670
DB 181 AAGATTGAAGCAGGAGGAGAGGGGACCAACAAAGATGAGATGGTGGATGCCATCACC 240
QY 2671 GACTCGATGAGCTGATCTGGGTGAATCTCTGGAGTTCCTGGATGAGTGGTGGAGGAGCTGT 2730
DB 241 GATTCAATGGACATGATTTGGGTAAACTCTGGAGCTGGAGATGGACAGGAGGAGCTGG 300
QY 2731 CTTGGCGGAGTTTCAATGGGTCAACAAAGATGTGACACGACTGAGCAACTGAACCTGA 2789
DB 301 CGTGCTACAGTCCATGAGTGGC--AGGGCCGGACATGACTGAGAGTGAACCTGAAT 357

RESULT 45

US-09-960-352-11719/c
; Sequence 11719, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11719
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 50-LIB3058-058-Q1-K1-E6
US-09-960-352-11719

Query Match 6.9%; Score 246.3; DB 1; Length 430;
Best Local Similarity 76.3%; Pred. No. 1.4e+02;
Matches 326; Conservative 0; Mismatches 92; Indels 9; Gaps 2;
QY 2360 CAGTGGTTCATGTGAGATGAGAGTGGACTGTGAAGAAAGCTGAGCAGCTGAAGAAATTG 2419
DB |||||
QY 430 CAGTAGTCATGTACCTAATGTCAGAAATGGACCAATGAAGAGGCCGATGGTTGCAGGATTG 371
DB |||||
QY 2420 ATGCTTTGAACCTGTGGTGTGGAGAAACATCTTTGAGAGTCCCTTGGACTGCAAGGAGAT 2479
DB |||||
QY 370 ATGCTTTCAAACCTGTGGTGTGGAGAAACATCTTTGAGAGTCCCTTGGACTGCAAGGACAT 311
DB |||||
QY 2480 CCAACCATGTCATCTGAAGGAGATCAGCCCTGGGATTTCTTTGGAGGAATGATGCTAA 2539
DB |||||
QY 310 CAAACCAAGTAATCTTAAGGAATCAACCTGAATCTTATTGGAAGGACTGATATAA 251
DB |||||
QY 2540 AGCTGAAACTCCAGTACTTTGGCCACCTGATCAGAAGAGCTGACTCA---CTGGAAGA 2596
DB |||||
QY 250 AG-----CTCTAATAGTTGGCCACCTGATGCCAGAGCTGACTCATTATGAAATGA 197
DB |||||
QY 2597 CCCTGATGCTGGAGGGATTGGGGCAGGAGGAGAGAGGGACACAGAGGATGAGATGCC 2656
DB |||||
QY 196 CCCTGATGCTGGGAAGACTGAAGTAAAGGAGAAAGGAGTGAAGAGGATGAGGTGAT 137
DB |||||
QY 2657 TGGATGGCATCACTGACTCGATGGACCTGAGTCTGGGTGAACCTCTGGAGTTGGTGATGG 2716
DB |||||
QY 136 TAAATGGCATCATCACTCAATGACCTGAGTGTGACCAACTCTGGAGCTGGTGAGA 77
DB |||||
QY 2717 ACAGGAGGCTGTCTTCCTGGCGGATTCATGGGGTCAAAAGAGTTGGAACAGCTGAGCA 2776
DB |||||
QY 76 ACAGGAGGCTGTCTTCCTGGCGGATTCATGGGGTCAAAAGAGTTGGAACAGCTGAGCTTAGTT 17
DB |||||
QY 2777 ACTGAAC 2783
DB |||||
16 ACTGAAC 10

Search completed: August 24, 2004, 13:39:08
Job time : 285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 13:40:03 ; Search time 73 Seconds
(without alignments)
3.947 Million cell updates/sec

Title: US-10-664-775-2

Perfect score: 3572

Sequence: 1 gtcaggagggcgagctga.....gcaacacagcagaagctt 3572

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 0.5

Searched: 45 seqs, 40327 residues

Total number of hits satisfying chosen parameters: 90

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : rst2.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	829.8	23.2	909	1	CC583446
C 2	824.6	23.1	926	1	ACCESSION:CC583446
C 3	824.3	23.1	949	1	CC479701
C 4	821.9	23.0	927	1	ACCESSION:CC479360
C 5	818.2	22.9	894	1	ACCESSION:CC508172
C 6	810.2	22.7	919	1	ACCESSION:CC544978
C 7	806.8	22.6	882	1	ACCESSION:CC577078
C 8	801.7	22.4	866	1	ACCESSION:CC514626
C 9	800.3	22.4	887	1	ACCESSION:CC586882
C 10	800.1	22.4	951	1	ACCESSION:CC520949
C 11	798.2	22.3	936	1	ACCESSION:CC548437
C 12	798.2	22.3	938	1	ACCESSION:CC594700
C 13	797.8	22.3	922	1	ACCESSION:B2882580
C 14	797.2	22.3	899	1	ACCESSION:CC480115
C 15	796.9	22.3	958	1	ACCESSION:CC475230
C 16	796.001	22.3	944	1	ACCESSION:CC593631
C 17	793.1	22.2	914	1	ACCESSION:CC503865
C 18	790.8	22.1	904	1	ACCESSION:CC536522
C 19	790.2	22.1	897	1	ACCESSION:CC476492
C 20	785.2	22.0	881	1	ACCESSION:CC579657
C 21	784.6	22.0	925	1	ACCESSION:CC500971
C 22	784.2	22.0	934	1	ACCESSION:CC543667
C 23	783.8	22.0	907	1	ACCESSION:CC547247
C 24	782.8	21.9	907	1	ACCESSION:CC480290
C 25	779.601	21.8	949	1	ACCESSION:CC589137
C 26	779.1	21.8	890	1	ACCESSION:CC541222
C 27	778.3	21.8	881	1	ACCESSION:CC546101
C 28	778.1	21.8	945	1	ACCESSION:CC511457
C 29	775.4	21.7	833	1	ACCESSION:CC495104
C 30	774.7	21.7	887	1	ACCESSION:CC539858
C 31	774.3	21.7	896	1	ACCESSION:CC528566
C 32	774.2	21.7	827	1	ACCESSION:CC508740
C 33	774.2	21.7	898	1	ACCESSION:CC512039

Query Match

23.2%; Score 829.8; DB 1; Length 909;

ALIGNMENTS

RESULT 1
CC583446/c

LOCUS

DEFINITION

genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

34 773.7 21.7 881 1 CC581072
35 771.3 21.6 915 1 CC536979
36 770.7 21.6 846 1 CC505149
37 769.501 21.5 947 1 CC591633
38 768.8 21.5 858 1 CC586052
39 768.5 21.5 877 1 CC529458
40 768.4 21.5 841 1 CC548709
41 768.4 21.5 865 1 CC533944
42 767.3 21.5 822 1 CC581625
43 767 21.5 861 1 CC595021
44 766.4 21.5 870 1 CC512505
45 765.9 21.4 862 1 CC540150

CC583446 909 bp DNA linear GSS 18-JUN-2003
CH240_379F21.T7 CHORI-240 Bos taurus genomic clone CH240_379F21,
genomic survey sequence.
CC583446
GSS.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 909)
Kolt,R., Scott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
Tsai,M., Cloutier,A., Lee,D., Ginn,N., Olson,T., Mayo,M., Chiu,R.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
Dalrymple,B.P. and Tellam,R.
Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other_GSSs: CH240_379F21.TABAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-8276
Email: rholt@bccgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 379 row: F column: 21
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..909
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="Breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_379F21"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/notes="Vector: pTABAC1.3; Site: 1: Mb01; Site 2: Mb01;
Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match

23.2%; Score 829.8; DB 1; Length 909;

Best Local Similarity 95.3%; Pred. No. 0.11;
Matches 855; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 533 GATGCTGTAAGAGCAATATTCATAGGACCTGGAATGTGAGGTCCTATGATCAAGCA 592
Db 897 GATGATGTAAGAGCAATATTCATAGGAACTGGAATGTGAGGTCCTATGATCAAGCA 838
QY 593 AATTGGAAGTGTCAAAACAAGAGATGGAAGATGAATGCAACAATCTTAGGAATCAGCG 652
Db 837 AATTGGAAGTGTCAAAACAAGAGATGGAAGATGAATGCAACAATCTTAGGAATCAGCG 778
QY 653 AACTAAATGGACTCGAATGGGTGAATTAATCTCAGATGACCATATATATCTACTACTGCG 712
Db 777 AACTGAAATGGACTCGAATGGGTGAATTAATCTCAGATGACCATATATATCTACTACTTCA 718
QY 713 GGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGTTCAACAAGAGTCGGAATG 772
Db 717 GGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGTTCAACAAGAGTCGGAATG 658
QY 773 CAGTACTTGGATGCACTCTCAAAACGACAGAGATGATCTCTGTTGTTTCCAAAGGCAAC 832
Db 657 CAGTACTTGGATGCACTCTCAAAATGACAGATGATCTCTGTTTCCAAAGGCAAC 598
QY 833 CATTCATATCACAGTAATCAAGTCTATGCCCCAACAGTAATGCTTGAAGAAGCTGAAG 892
Db 597 CATTCATATCACAGTAATCAAGTCTATGCCCCAACAGTAATGCTTGAAGAAGCTGAAG 538
QY 893 TTGAACGCTCTATGAAGACCTCAAGACCTTTAGAACTAAACACCCCAAAAAGATGCTCC 952
Db 537 TTGAACGCTCTATGAAGACCTCAAGACCTTTAGAACTAAACACCCCAAAAAGATGCTTC 478
QY 953 TTCTCATATAGGGAATGGAATGCAAAAGTAGGAAGCAAAACACCTGGAGTAACAG 1012
Db 477 TTTTCATATAGGGAATGGAATGCAAAAGTAGGAAGTAAGAAACACCTGGAGTAACAG 418
QY 1013 GCAAAATTTGGCTTGGATACGGAATGGAAGCAAGGCAAGACTAATAGATTTTGCAAG 1072
Db 417 GCAAAATTTGGCTTGGATACGGAATGGAAGCAAGGCAAGACTAATAGATTTTGCAAG 358
QY 1073 AAAATGCACTGGTCATAGCAACACCTCTTCCAAACCAACAGAGAAAGACTCTACACAT 1132
Db 357 AAAATGCACTGGTCATACAAACACCTCTTCCAAACCAACAGAGAAAGACTCTACACAT 298
QY 1133 GGAATCACAGATGCTCAACACCGAATCAGATGATTAATATTTCTTTGAGGCAAGAT 1192
Db 297 GGAATCACAGATGCTCAACACCGAATCAGATGATTAATATTTCTTTGAGGCAAGAT 238
QY 1193 GGAGAACTCTATACAGTCAAGCAAAACACAGACGAGCTTACTGTGGCTCAGATCATG 1252
Db 237 GGAGAACTCTATACAGTCAAGCAAAACACAGACGAGCTTACTGTGGCTCAGATCATG 178
QY 1253 AACTCCTTTATGCAAAATTCAGATTAATTAAGAAAGTAGGAAACCACTAGATCAC 1312
Db 177 AACTCCTTTATGCAAAATTCAGATTAATTAAGAAAGTAGGAAACCACTAGATCAC 118
QY 1313 TCAGTAAGACCTAAATCCATCCCTTATGATTTATACAGTGAAGTGAAGATAGATTTA 1372
Db 117 TCAGTAAGACCTAAATCCATCCCTTATGATTTATACAGTGAAGTGAAGATAGATTTA 58
QY 1373 AGGCCCTAGATCTGATAGACAGATGCTAATGAACCTATGACAGAGGTTTCATGACA 1429
Db 57 AGGCCCTAGATCTGATAGATGAGTGCCTGATGAACCTATGGAATGAGGTTTCGTGACA 1

RESULT 2
LOCUS CC479701
DEFINITION CH240_306H6.T7 CHORI-240 Bos taurus genomic clone CH240_306H6,
genomic survey sequence.
ACCESSION CC479701
VERSION CC479701.1 GI:31757509
KEYWORDS GSS.
SOURCE Bos taurus (cow)

ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactylia; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 926)

AUTHORS

Holt, R., Stott, J., Yang, G., Barber, S., Smalish, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M., Chiu, R.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Quin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
Dalrymple, B. P. and Tellam, R.

TITLE

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

JOURNAL

Unpublished (2003)

COMMENT

Other GSSs: CH240_306H6.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bccg.ca

Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/orderinginformation.htm). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 306 row: H column: 6

Seq primer: T7

Class: BAC ends.

FEATURES

Location/Qualifiers

1..926

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="bred: Hereford"

/db_xref="taxon:9913"

/clone="CH240_306H6"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

Query Match

Best Local Similarity 23.1%; Score 824.6; DB 1; Length 926;

Matches 875; Conservative 0; Mismatches 50; Indels 2; Gaps 2;

QY 45 GGTAAAGGAGCAGTAGCTCGCTTTGCTGGAGCAGCGTAAAGAGATACCCACGCCCAAG 104

Db 1 GGTAAAGGAGCAGTAGCTCGCTTTGCTGGAGCAGCGTAAAGAGATACCTCACGCCCAAG 60

QY 105 GTAACAGAAACCCAGTAGTGGTAGGTGTGTGAGAGGCATCAGAGGCAGCATAC 164

Db 61 GTAACAGATACCCAGTAGTGGTAGGTGTGTGAGAGGCATCAGAGGCAGCATAC 120

QY 165 TGAAACCATACACGAGAAACTAGTCAATCAATCACTAGGACCAAGCCCTTTGCTA 224

Db 121 TGAAACCATACCTCAGAGAAACTAGTCAATCAATCACTAGGACCAAGCCCTTTGCTA 180

QY 225 ACTCAATCAACTAAGCCATGCCGTGGGCAACCCAGATGGGAGGTCTAGTGAGA 284

Db 181 ACTCAATCAACTAAGCCATGCCGTGGGCAACCCAGATGGGAGGTCTAGTGAGA 240

QY 285 GATCTGACAGAAATGTGGTCCACTGCGAGAGGGAAT-GCAACCACTTCAGTATTCCTGCC 343

Db 241 GATCTGACAGAAATGTGGTCCACTGCGAGAGGGAATGGAACCACTTCAGTATTCCTGCC 300

QY 344 TTGAGAACCCCATGACAGTATGAAAAGGCAAAATGATAGTACTGAAAGAGAACTCC 403

Db 301 TTGAGAACCCCATGACAGTATGAAAAGGCAAAATGATAGTACTGAAAGAGAACTCC 360

(<http://www.chori.org/bacpac/bovine240.htm>). For BAC library availability, please contact Pieter de Jong (pdejong@mail.choi.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada.

Plate: 306 row: K column: 22

Seq primer: 17

Class: BAC ends.

Location/Qualifiers

source

1..949
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_306K22"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

Query Match 23.1%; Score 824.3; DB 1; Length 949;
Best Local Similarity 94.2%; Pred. No. 0.11;
Matches 896; Conservative 0; Mismatches 48; Indels 7; Gaps 4;

1667 GAAAAGCAGAGAGTTCCAGAAACACATCTATTCTGCTTTTATTGACTATCCAAAAGCC 1726
Db 3 GGAAGCAGAGAGTTAAAGAAACACATCTATTCTGCTTTTATTGACTATCCAAAAGCC 62
1727 TTTGACTGTGGGGTCCAAATAAATCTGGAATAATCTGAAAGGGATGGAATACAGAC 1786
Db 63 TTTGACTGTGGATCCAAATAAATCTGGAATAATCTGGAATAATCCAGAC 120
1787 CACCTGACCTGCTCTTGAATAATTTGATGAGTCCAGGTCCAGGACCAACAGTTAGAACTGA 1846
Db 121 CACCTGACCTGCTCTTGAATAATTTGATGAGTCCAGGTCCAGGACCAACAGTTAGAACTGA 180
1847 CATGGAACCAACAGACTGTTTCAAGTAGGAAGAAGATGTCAAGGCTGTATATTGTCA 1906
Db 181 CATGGAACCAACAGACTGTTTCAAGTAGGAAGAAGATGTCAAGGCTGTATATTGTCA 240
1907 CCGGCTTTTAACTTCTATGACAGAG-ACATCATGAGAAACGCTGGCTGGAAGAGCA 1965
Db 241 CCGGCTTTTAACTTCTATGACAGAGTACATAATGAGAAACGCTGGACTGGAAGAGCA 300
1966 CAGCTGGAATCAAGATTCGCGGAGAAATAGCAATAACCTCAGATATCAGATGATACC 2025
Db 301 CAGCTGGAATCAAGATTCGCGGAGAAATAGCAATAACCTCAGATATCAGATGATACC 360
2026 ACCCTTATGCGAGAAAGTGAAGAGGAACCTTAAAGAGCTCTTGTATGAAGTGAAGAGGAG 2085
Db 361 ACCCTTATGCGAGAAAGTGAAGAGGAACCTTAAAGAGCTCTTGTATGAAGTGAAGAGGAG 420
2086 AGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACCAAGATCATGCACTCTGCTCC 2145
Db 421 AGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAACCAAGATCATGCACTCTGCTCC 480
2146 ATCACTTCAATGGAAATAGATGGGAAACAGATGGAACACAGTGTCAAGCTTTATTTTGG 2205
Db 481 ATCACTTCAATGGAAATAGATGGGAAACAGATGGAACACAGTGTCAAGCTTTATTTT 538
2206 GGGGCTCCAAATCATCTGACAGATGGTGTGACTGAGCCCATGAATTAAGACACTTACTCC 2265
Db 539 GGGGCTCCAAATCATCTGACAGATGGTGTGACTGAGCCCATGAATTAAGACACTTACTCC 598
2266 TTGGAAGAAAGTTA--ACCACTAGATAGCATATTGAAAGACAGAGACATTACCTTGC 2323
Db 599 TTGGAAGAAAGTTAAGCACTAGATAGCATATTGAAAGACAGAGACATTACCTTGT 658
2324 CAACAAGGCCCATCTAGTCAAGGCTATGTTTTTCCAGTGGTCAATGTATGGATGTGAGA 2383

404 CCAGGTCAAGTGGTCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGA 463
Db 361 CCAGGTCAAGTGGTCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGA 420
464 ATGAAGAGATGGAGCCAAAGCAAAAGAAATACCCAGCTGGAGTGGACTGGTGATATAA 523
Db 421 AAGAGGATGGAGCCAAAGCAAAAGAAATACCCAGCTGGAGTGGACTGGTGATAGAA 480
524 GCAAGGTCGATGCTGTAAGAGCAATATTGCTATAGGAACCTGGAATGTCCAGTCCATGA 583
Db 481 GTRAGGTCCAATGCTGTAAGAGCAATATTGCTATAGGAACCTGGAATGTCCAGTCCATGA 540
584 ATCAAGCAAAATTTGGAAGTGGTCAAAACAGAGATGGCAAGAGTGAATGCAACATCTAG 643
Db 541 ATCAAGCAAAATTTGGAAGTGGTCAAAACAGAGATGGCAAGAGTGAATGCAACATCTAG 600
644 GAATCAGCGAACTAAATGACTGGAATGGTGAATTTAACTCAGATGACCATATATCT 703
Db 601 GAATCAGCGAACTAAATGACTGGAATGGTGAATTTAACTCAGATGACCATATATCT 660
704 ACTACTCGGGCAGGAATCCCTCAGAAGAAATGGAGTAGCCATCATGCTCAACAAAGAG 763
Db 661 ACTACTCGGGCAGGAATCCCTCAGAAGAAATGGAGTAGCCATCATGCTCAACAAAGAG 720
764 TCCGAATGAGTACTTGGATGAGTCTCAAAACAGAGATGATCTCTGTTTGTTC 823
Db 721 ACCGAATGAGTACTTGGATGAGTCTCAAAACAGAGATGATCTCTGTTTGTTC 780
824 AAGCAAAACCATTCATATCAGATTAATCCAGTCTATGCCCCCAACCAAGTGAAG 883
Db 781 AAGCAAAACCATTCATATCAGATTAATCANAGTCTATGCCCAACCAAGTGAAG 840
884 AAGCTGAAGTTGACGCTCTTATGAGACCTTACAGACCTTTTAGACTACACCCAAA 943
Db 841 AAGCTGAAGTTGACGCTCTTATGAGACCTTACAGACCTTTTAGACTACACCCAAA 899
944 AAGATGCTCTCTCATTTATAGGGACT 970
Db 900 AAGATGCTCTCTCATTTATAGGGAAAT 926

CC479360 949 bp DNA linear GSS 16-JUN-2003
CH240_306K22.T7 CHORI-240 Bos taurus genomic clone CH240_306K22,
genomic survey sequence.

CC479360

CC479360.1 GI:31756840

GSS.

Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.

1 (bases 1 to 949)

Holt, R., Stott, J., Yang, G., Barber, S., Smal, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Gira, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Matheson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
Dalrymple, B.P. and Tellam, R.

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)

Other GSSs: CH240_306K22.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bccsc.ca

Clones are derived from the bovine BAC library CHORI-240

```

Db      659 CAACAAAGGTTTCATCTAGTCAAGGCTATGGTTTTTCCCTGGTGCATGTATGGATGTGAGA 718
QY      2384 GTTGGACTGTGAAGAAAGCTGAGCACTGAAGAATTGATGCTTTTGAACCTGTGGTTGGA 2443
Db      719 TTTGGACTGTGAAGAAAGCTGAGCACTGAAGAATTGATGCTTTTGAACCTGTGGTTGGA 778
QY      2444 GAAGACTCTTGAAGTCCCTTGGACTGCAAGAGATCCAAACAGTCCATTTCTGAAGGAGA 2503
Db      779 GAAGACTCTTGAAGTCCCTTGGACTGCAAGAGATCCAAACAGTCCATTTCTGAAGGAGA 838
QY      2504 TCAGCCCTGGGATTTCTTTGGAGAAATGATCTTAAGCTGAAGCTGAAGCTGAAGCTGAAGCT 2563
Db      839 TCAGCCCTGGGATTTCTTTGGAGAAATGATCTTAAGCTGAAGCTGAAGCTGAAGCTGAAGCT 898
QY      2564 ACCTGATCAGAGAGCTGACTCACTGGAAAGACCCCTGATGCTGGGAGGGA 2614
Db      899 ACCTCATGCAAGAGTGTGACTCATTTGGAAAGAGACTCTGATGCTGGGAGGGA 949

RESULT 4
CC508172/c
LOCUS   CH240_350K18.T7 CHORI-240 Bos taurus genomic clone CH240_350K18,
DEFINITION genomic survey sequence.
ACCESSION CC508172
VERSION   CC508172.1 GI:31826465
KEYWORDS  GSS.
SOURCE    Bos taurus (cow)
ORGANISM  Bos taurus

REFERENCE
AUTHORS  Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
Dairymple,B.P. and Tellam,R.
TITLE    Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
JOURNAL  Unpublished (2003)
COMMENT  Other GSSs: CH240_350K18.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W.10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 350 row: X column: 18
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1. .927
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clones="CH240_350K18"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC

```

library (Male) produced by Pieter de Jong

Query Match 23.0%; Score 821.9; DB 1; Length 927;

Best Local Similarity 93.8%; Pred. No. 0.11; Indels 7; Gaps 2;

Matches 876; Conservative 0; Mismatches 51;

```

QY      408 GTCAGTAGTGGCCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAATCA 467
Db      927 GTCAGTAGTGGCCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGAATCA 868
QY      468 AGAGATGGAGCCCAACCAAAAGAAATACCAGCTGTGGATGTGACTGGTGAATAAGCA 527
Db      867 AGGAATGGAGCCCAACCAAAAGAAATACCAGCTGTGGATGTGACTGGTGAATAAGCA 808
QY      528 GGTCCGATGCTGTAAAGAGCAATATTCATAGGAACCTGGAATGTGAGTCCATGAATCA 587
Db      807 GGTCCGACGCTGTAAAGAGCAATATTCATAGGAACCTGGAATGTGAGTCCATGAATCA 748
QY      588 AGSCAAATTTGGAAGTGGTCAAAACAAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAAT 647
Db      747 AGSCAAATTTGGAAGTGGTCAAAACAAGAGATGGCAAGAGTGAATGTCAACATTTCTAGGAAT 688
QY      648 CAGCGAATCTAAATGAGTGGAAATGGTGAATTTAACTCAGATGACCATTTATATCTACTA 707
Db      687 CAGCGAATCTAAATGAGTGGAAATGGTGAATTTAACTCAGTGGACCCCTTATATCTGCTA 628
QY      708 CTGCGGGCAGGAATCCCTCAGAAAGAAATGGAGTAGCCATTCATGGTCAACAAAGAGTCGG 767
Db      627 CTGTCGGCAGGAATCCCTCAGAAAGAAATGGAGTAGCCGTCATGGTCAACAAAGAGTCGTG 568
QY      768 AAATGAGTACTGGTGGTGCAGTCTCAAAAACGACAGATGATCTCTGTTGTTTCCAAAGG 827
Db      567 AAATGAGTACTGGTGGTGCAGTCTCAAAAACGACAGATGATCTCTGTTGTTTCCAAAGG 508
QY      828 CAAACCATTTCAATATCACAGTAATCCAAGTCTCTATCCCCCAACCAAGTAATGTCTGAAGAAGC 887
Db      507 CAAACCATTTCAATATCACAGTAATCCAAGTCTCTATGCCCAACCAAGTAATGTCTGAAGAAGC 448
QY      888 TGAAGTTGAGCGGTCTTATGAAGACTACAGACCTTTTAGACTAACACCCCAAAAAGA 947
Db      447 -----TGAACGGTTCTTACGAGAGACCTTACAGGCGCTTTTAGAACCTTAAACCCCAAAAAGA 394
QY      948 TGTCTCTTCTCATTTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACAACCTGGAGT 1007
Db      393 TGTCTCTTCTCATTTATAGGGGACT-GAGTGCAGAAAGTAGGAATCAAGAAACACCTGGAGT 335
QY      1008 AACAGGCAAAATTTGGCCTTGGAAATACGGAATGAAGCAGGCAAGAGCTAATAGAGTTTTG 1067
Db      334 AACAGGCAAAATTTGGCCTTGGAAATACGGAATGAAGCAGGCAAGAGCTAATAGAGTTTTG 275
QY      1068 CCAAGAAATGCACTGGTTCATAGCAAAAACACCTCTTCCAAACAACAAGAGAGAGCTCTA 1127
Db      274 CCAAGAAATGCACTGGTTCATAGCAAAAACACCTCTTCCAAACAACAAGAGAGAGCTCTA 215
QY      1128 CACATGGACATCACAGATGGTCAACACCGAAATCAGATTGATTATATTTTTCAGGCA 1187
Db      214 CACATGGACATCAACAGATGGTCCACACCAAAATCAGACTGATTATATTTTTCAGGCA 155
QY      1188 TAGATGGAGAGCTCTATACAGTCAAGCAAAACAGACCGAGGCTTACTGTGCTCAGA 1247
Db      154 TAGATGGAGAGCTCTATACAGTCAAGCAAAACAGACCGAGGCTTACTGTGCTCAGA 95
QY      1248 TCATGAATCTCTTATTTGCCAAATTCAGACTTTAAATTTGAAGAAAGTAGGGAAAACCTCTAG 1307
Db      94 TCATGAATCTCTTATTTGCCAAATTCAGACTTTAAATTTGAAGAAAGTAGGGAAAACCTCTAG 35
QY      1308 ATCACTCAGGTAAAGACCTTAAATTCNAATCCCTTAT 1341
Db      34 GCCATTTCAGGTATGACCTTAAATTCNAATCCCTTAT 1

```

RESULT 5
CC545978

LOCUS CC544978 894 bp DNA linear GSS 17-JUN-2003
DEFINITION CH240_427A14.TARBAC13P2 CHORI-240 Bos taurus genomic clone
ACCESSION CH240_427A14, genomic survey sequence.
VERSION CC544978
KEYWORDS CC544978.1 GI:31863262
SOURCE GSS.
ORGANISM Bos taurus (cow)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 894)
Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., Keele, J.W. and Kappes, S.M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)
Other GSSs: CH240_427A14.77
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 427 Row: A Column: 14
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..894
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="bred: Hereford"
/db_xref="taxon:9913"
/clone="CH240_427A14"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull LI Domino 99375; CHORI-240 Bovine BAC
Library (Male) produced by Pieter de Jong"
Query Match 22.9%; Score 818.2; DB 1; Length 894;
Best Local Similarity 96.0%; Pred. No. 0.12;
Matches 849; Conservative 0; Mismatches 33; Indels 2; Gaps 1;
QY 298 GTGTCCTCACTGAGAGGAAT--GCAACCACTTCAGTATCTTGCCTTGAGAACCCCA 355
DB 11 GTGTCCTCACTGAGAGGAATGTCGCAACCACTTCAGTATCTTGCCTTGAGAACCCCA 70
QY 356 TGAACAGATGAAAAGGCAAAATGATAGGATCTGAAAGAGCACTCCCGAGTCAGTAG 415
DB 71 TGAACAGATGAAAAGGCAAAATGATAGGATCTGAAAGAGCACTCCCGAGTCAGTAG 130
QY 416 GTGCCCCATATGCTACTCTGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGATGG 475
DB 131 GTGCCAATATGCTACTGAGATCAGTGGAGAAATACTCCAGAAAGATGAAGATGG 190
QY 476 AGCCAAAGCAAAAGAAATACCCAGCTGGGATGTCAGTGGTATAGCATAGCAAGTCCGAT 535
DB 191 AGCCAAAGCAAAAGAAATACCCAGCTGGGATGTCAGTGGTATAGCATAGCAAGTCCGAT 250
QY 536 GCTGTAAAGAGCAATATTTCATAGGACCTCGAATGTCCAGTCCATGAATCAAGGCAAT 595

DB 251 GCTGTAAAGAGCAATATTTCATAGGAACTGGATGTCAGGTCCATGAATCAAGGCAAT 310
QY 596 TGAAGTGTCAAACAAGAGATGGCAAGAGTGAATGTCAACATTTAGGAATCAGGCAAC 655
DB 311 TGAAGTGTCAAACAAGAGATGGCAAGAGTGAATGTCAACATTTAGGAATCAGGCAAC 370
QY 656 TAAATGGTGAATGGGTGAATTTAACTCAGATGACCATATATCTACTACTCTGCGGC 715
DB 371 TGAATGGTGAATGGGTGAATTTAACTCAGATGACCATATATCTACTACTCTGCGGC 430
QY 716 AGGAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAGAGTCCGAAATGCGAG 775
DB 431 AGGAATCCCTCAGAAGAAATGGAGTAGCCATCATGTCAACAAAGAGTCCGAAATGCGAG 490
QY 776 TACTTGGTGCAGTCTCAAACACGACAGATGATCTCTGTTTGTTCCTCAAGGCAACCAT 835
DB 491 TACTTGGTGCAGTCTCAAACACGACAGATGATCTCTGTTTGTTCCTCAAGGCAACCAT 550
QY 836 TCAATATCAGATTAATCCAAAGTCTATGCCCCAACCAAGTAATGCTGAAGAGCTGAAGTTG 895
DB 551 TCAATATCAGATTAATCCAAAGTCTATGCCCCAACCAAGTAATGCTGAAGAGCTGAAGTTG 610
QY 896 AAGGTCTCTATGAGACCTTACAGACCTTTTAGAAGTAAACCCCAAAAGATGCTCTTC 955
DB 611 AAGGTCTCTATTAAGACCTTACAGACCTTTTAGAAGTAAACCCCAAAAGATGCTCTTC 670
QY 956 TCATTATAGGGACTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGCGAGTAACAGGCA 1015
DB 671 TCATTATAGGGACTGGAATGCAAAAGTAGGAAGTCAAGAAACACCTGCGAGTAACAGGCA 730
QY 1016 AATTGGCTTGGAAATACGGAATGAAAGAGGGCAAGAGATAATAGAGTTTGGCAAGAA 1075
DB 731 AATTGGCTTGGAAATACGGAATGAAAGAGGGCAAGAGATAATAGAGTTTGGCAAGAA 790
QY 1076 ATGCACTGTCTAGCAACACCCCTCTCCCAACACAGAGAGACACTCTACACATGA 1135
DB 791 ATGCACTGTCTAGCAACACCCCTCTCCCAACACAGAGAGACTCTATACATGA 850
QY 1136 CATCACCAGATGCTCAACACCCGAAATCAGATGATTAATTTCTT 1179
DB 851 CATCACCAGATGCTCAACACCCGAAATCAGATGATTAATTTCTT 894

RESULT 6
CC577078/c
LOCUS
DEFINITION CH240_455012.TARBAC13P2 CHORI-240 Bos taurus genomic clone
ACCESSION CH240_455012, genomic survey sequence.
VERSION CC577078
KEYWORDS CC577078.1 GI:31920393
SOURCE GSS.
ORGANISM Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 919)
Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., Keele, J.W. and Kappes, S.M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)
Other GSSs: CH240_455012.77
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca

Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

Query Match 22.6%; Score 806.8; DB 1; Length 882;
 Best Local Similarity 94.7%; Pred. No. 0.14;
 Matches 835; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 476 AGCCAAAGCAAAAGAAATACCAGCTGTGGATGTGACTGTGGTATAGCAAGAGTCCGAT 535
 DB 1 AGCCAAAGCAAAAGCAATACCTGTGGATGTGCTGTGTATAGCAAGAGTCCGAT 60

QY 536 GCTGTAAAGACATATTTGCTAGAGCACTGGATGTGCTAGTCCATGAATCAAGGCAAT 595
 DB 61 GCTGTAAAGACATATTTGCTAGAGCACTGGATGTGCTAGTCCATGAATCAAGGCAAT 120

QY 596 TGGAAAGTGGTCAAAACAGAGATGGCAAGAGTGAATGTCAACATTTAGGAATCAGCAAC 655
 DB 121 TGGAAAGTGGTCAAAACAGAGATGGCAAGAGTGAATGTGACATTTCTAGGAATCAGCAAC 180

QY 656 TAAATGGACATGGATGGTGAATTTAACTCAGATGACCATATATCTACTTACTTGGGGC 715
 DB 181 TGAATGGACATGGATGGTGAATTTAACTCAGATGACCATATATCTACTTACTTGGGGC 240

QY 716 AGGAATCCCTCAGAGAAATGGATGGATCATGCTCAACAAAGAGTCCGAATCGAG 775
 DB 241 AGGAATCCCTCAGAGAAATGGATGGATCATGCTCAACAAAGAGTCCGAATCGAG 300

QY 776 TACTTGGATGAGTCTCAAAACGACAGATGATCTCTGTGTTTCCAGGCAAAACAT 835
 DB 301 TGCITGGATGCAATCTCAAAACGACAGATGATCTCTGTGTTTCCAGGCAAAACAT 360

QY 836 TCAATATCAGAGTATCCAGTCTATGCCCCAAGCAGTATGCTGAAGAGTCAAGTTG 895
 DB 361 TCAATATCAGAGTATCCAGTCTATGCCCCAAGCAGTATGCTGAAGAGTCAAGTTG 420

QY 896 AACGCTCTATGAGACCTCAAGACCTTTTAGAATTAACACCCCAAAAGATGCTCTTC 955
 DB 421 AACGCTCTATGAGACCTCAAGACCTTTTAGAATTAACACCCCAAAAGATGCTCTTC 480

QY 956 TCATTTAGGGAGTGGAAATCAAAAGTAGAAGCAAGAAACCTGGAGTAAACAGCA 1015
 DB 481 TCATTTAGGGAGTGGAAATCAAAAGTAGAAGTCAAGAAACCTGGAGTAAACAGCA 540

QY 1016 AATTGGCTTTGGATATCGGATGAAGCAGGCAAGCACTAATAGATTTTGGCAAGAA 1075
 DB 541 AATTGGCTTTGGATATCGGATGAAGCAGGCAAGCACTAATAGATTTTGGCAAGAA 600

QY 1076 ATGCACTGGTATAGCAAAACACCTCTTCCAAACACAAAGAGAGTCTTACATGGA 1135
 DB 601 ATGCACTGGTATAGCAAAACACCTCTTCCAAACACAAAGAGAGTCTTACATGGA 660

QY 1136 CATCAGAGATGGTCAACACCAATCAGATGATATATCTTGTGAGCAAGATGGA 1195
 DB 661 CATCAGAGATGGTCAACACCAATCAGATGATATATCTTGTGAGCAAGATGGA 720

QY 1196 GAAGCTCTATACAGTCAAGCAAAACACAGCAGAGCTTACTGTGGCTCAGATCATGAAC 1255
 DB 721 GAAGCTCTATACAGTCAAGCAAAACACAGCAGAGCTTACTGTGGCTCAGATCATGAAC 780

QY 1256 TCCTTATTTGCCAAATTCAGCTTAATTTGAAGAAAGTAGGGAACCACTAGATCACTCA 1315
 DB 781 TCCTTATTTGCCAAATTCAGCTTAATTTGAAGAAAGTAGGGAACCACTAGATCACTCA 840

QY 1316 GGTAGACCTTAATCCATCCCTTATGATTTATACAGTGAAG 1357
 DB 841 GGTAGACCTTAATCCATCCCTTATGATTTATACAGTGAAG 882

RESULT 8
 CC586882/c
 LOCUS
 DEFINITION CH240_384H21.T7 CHORI-240 Bos taurus genomic clone CH240_384H21, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

CC586882
 CC586882.1 GI:31939990
 GSS.
 Bos taurus (cow)
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 866)
 Holt, R., Cloutier, J., Yang, G., Barber, S., Smailus, D., Prabhu, A., L.,
 Tsai, M., Stott, J., Lee, D., Girn, N., Olson, T., Mayo, M., Chiu, R.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Quin, R., Chan, A., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B. P. and Tellam, R.
 Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 Unpublished (2003)
 Other GSSs: CH240_384H21.TARBAC13P2
 Contact: Rob Holt
 Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bcgsc.ca
 Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.choi.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.html). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 384 row: H column: 21
 Seq primer: T7
 Class: BAC ends.

FEATURES
 source

Location/Qualifiers
 1..866
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_384H21"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /notes="Vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 22.4%; Score 801.7; DB 1; Length 866;
 Best Local Similarity 96.1%; Pred. No. 0.15; Indels 1; Gaps 1;
 Matches 832; Conservative 0; Mismatches 33;

QY 510 GACTGTGATATAGCAAGTCCGATGCTGTAAAGAGCAATATTGCATAGGAACCTGGAA 569
 DB 866 GACTGTGATAGCAAGTCCGATGCTGTAAAGAGCAATATTGCATAGGAACCTGGAA 807

QY 570 TGTCAAGTCCATCAATCAAGGCAATGGAAGTGTCAACAGAGATGCGAGGTGAA 629
 DB 806 TGTCAAGTCCATCAATCAAGGCAATGGAAGTGTCAACAGAGATGCGAGGTGAA 747

QY 630 TGTCAACATTTCTAGGAATCAGCAACTAAAATGGACTGGAATGGGTGAATTTAACTCAGA 689
 DB 746 TGTCAACATTTCTAGGAATCAGCAACTAAAATGGACTGGAATGGGTGAATTTAACTCAGA 687

QY 690 TGACCATATATCTACTTCTGCGGGCAGGAATCCCTCAGAAGAAATGGAGTAGCCATCAT 749
 DB 686 TGACCATATATCTACTTCTGCGGGCAGGAATCCCTCAGAAGAAATGGAGTAGCCATCAT 627

QY 750 GGTCAACAAAGAGTCCGAAATGCAGTACTTGTGATCGAGTCTCAAAAACGACAGAAATGAT 809
 DB 626 GGTCAACAAAGAGTCCGAAATGCAGTACTTGTGATCGAGTCTCAAAAACGACAGAAATGAT 567

QY 810 CTCCTGTTTGTTCCTCAAGCGCAACCACTTCAATATACAGTATCCAGTCTATGCCCAAC 869
 Db 566 CTCCTGTTTGTTCCTCAAGCGCAACCACTTCAATATACAGTATCCAGTCTATGCCCAAC 507
 QY 870 CAGTAATGCTGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACCTTACAGACCTTTTGA 929
 Db 506 CAGTAACACTGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACCTTACAGACCTTTTGA 447
 QY 930 ACTAACACCCCAAAAGAGATGCTTCTATATATAGGGAAGTGGATGCAAAAGTAGGAG 989
 Db 446 ACTAACACCCCAAAAGAGATGCTTCTTTCATTATA-GGGAAGTGGATGCAAAAGTAGGAG 388
 QY 990 CAAGAAGACCTGAGTAACAGGCAAAATTTGGCTTGGATACGGAATGAAGCAGGCA 1049
 Db 387 TCAAGGAATACCTGGATGAACAGGCAAAATTTGGCTTGGATACGGAATGAAGCAGGCA 328
 QY 1050 AAGACTAATAGAGTTTTCGCAAGAAATGCACTGCTGTCTATAGCAAAACCTCTTCCAACA 1109
 Db 327 AAGGCTAATAGAGTTTTCGCAAGAAATGCACTGCTGTCTATAGCAAAACCTCTTCCAACA 268
 QY 1110 ACACAAGAGAGACTCTACATGACATCACCAGATGCTCAACACCGAATCAGATTGA 1169
 Db 267 ACACAAGAGAGACTCTACATGACATCACCAGATGCTCAACACCGAATCAGATTGA 208
 QY 1170 TTATATCTTTTCGACCCAAAGATGAGAAGCTCTATACAGTCAGCAAAACCAAGCAGG 1229
 Db 207 TTATATCTTTTCGACCCAAAGATGAGAAGCTCTATACAGTCAGCAAAACCAAGCAGG 148
 QY 1230 AGCTTACTGCTCAGATCAGTACCTCTTATGCGCAATTCAGACTTAAATGAAGAA 1289
 Db 147 AGCTTACTGCTCAGATCAGTACCTCTTATGCGCAATTCAGACTTAAATGAAGAA 88
 QY 1290 AGTAGGGAAACCACTAGATCACTCAGGTAAAGACTTAAATCCATCCCTTATGATTATAC 1349
 Db 87 AGTAGGGAAACCACTAGATCACTCAGGTAAAGACTTAAATCCATCCCTTATGATTATAC 28
 QY 1350 AGTGAAGTGAATAGATTAAAG 1375
 Db 27 AGTGAAGTGAATAGATTAAAG 2
 RESULT 9
 CC520949
 LOCUS CH240_368M20.TARBA13P2 CHORI-240 Bos taurus genomic clone
 DEFINITION CH240_368M20, genomic survey sequence.
 ACCESSION CG520949
 VERSION CG520949.1 GI:31839237
 KEYWORDS GSS
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 887)
 AUTHORS Holt, R., Stott, J., Yang, G., Barber, S., Smalley, D., Prabhu, A.-L.,
 Tsai, M., Cloutier, A., Lee, D., Ginn, N., Olson, T., Mayo, M.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B.P., and Tellam, R.
 TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 JOURNAL Unpublished (2003)
 COMMENT Other_GSSs: CH240_368M20.77
 Contact: Rob Holt
 Sequencing The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bccsc.ca
 Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.choi.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 368 row: M column: 20
 Seq primer: SP6
 Class: BAC ends.

FEATURES source

Location/Qualifiers
 1..887
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_368M20"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 22.4%; Score 800.3; DB 1; Length 887;
 Best Local Similarity 94.6%; Pred. No. 0.14;
 Matches 839; Conservative 0; Mismatches 47; Indels 1; Gaps 1;
 QY 289 TGACAGAATGTGTCACCTGGAGAGAGGAAT-GCAAAACCACTTCAGTATCTTGCCTGA 347
 Db 1 TGACAGAATGTGTCACCTGGAGAGAGTAATGGCAACCACTTCAGTATCTTGCCTGA 60
 QY 348 GAACCCCATGAACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAATCTCCCAAG 407
 Db 61 GAACCCCATGAACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGAATCTCCCAAG 120
 QY 408 GTTCAGTAGTGGCCCATATGCTTCTGAGATCAGTGGGAATTAATCTCCAGAGAAGATGA 467
 Db 121 GTTCAGTAGTGGCCCATATGCTTCTGAGATCAGTGGGAATTAATCTCCAGAGAAGATGA 180
 QY 468 AGAGATGGAGCCAAAGCAAAAGATAATCCAGCTGTGGATGTGACTGGTGTATATAGCAA 527
 Db 181 AGGACGGAACCAAGCAAAACACACCCAGCTGTGGATGTGACTGGTGTATAGCAA 240
 QY 528 GGTCCGATCTGTAAAGCAATTTGCAATAGGAACCTGGAAATGTCAAGTCCATGAATCA 587
 Db 241 GGTCCGATCTGTAAAGCAATTTGCAATAGGAACCTGGAAATGTCAAGTCCATGAATCA 300
 QY 588 AGGCAAAATTTGGAAGTGGTCAAAACAGAGATGGCAAGATGAATGTCAACATCTTAGGAAT 647
 Db 301 AGGCAAAATTTGGAAGTGGTCAAAACAGAGATGGCAAGATGAATGTCAACATCTTAGGAAT 360
 QY 648 CAGCGAATCAAAATGGAATGGTGAATGGTGAATTTAACTCAGATGACCAATATATCTACTA 707
 Db 361 CAGCGAATCAAAATGGAATGGTGAATGGTGAATTTAACTCAGATGACCAATATATCTACTA 420
 QY 708 CTGCGGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGTGCAACAAAGAGTCCG 767
 Db 421 CTGCGGGCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGTGCAACAAAGAGTCCG 480
 QY 768 AAATGCACTACTTGGATGCACTCTCAAAACGACAGAATGATCTCTGTGTTCCTCAAGG 827
 Db 481 AAATGCACTACTTGGATGCACTCTCAAAACGACAGAATGATCTCTGTGTTCCTCAAGG 540
 QY 828 CAACCAATTCATATACAGTAAATCCAGTCTATGCCCAACCAAGTATGCTGAGAGGC 887
 Db 541 CAACCAATTCATATACAGTAAATCCAGTCTATGCCCAACCAAGTATGCTGAGAGGC 600
 QY 888 TGAAGTTGAACCGTCTCTATGAGAGCTTCAAGACCTTTTAGAACCTTAACACCAAAAAAGA 947
 Db 601 TGAAGTTGAATGATCTCTATGAGAGCTTCTATAGATCTTTTAGAACCTTAACACCAAAAAAGA 660
 QY 948 TGTCTCTCTCATTTAGGGGACTGGAATGCAAAAGTAGGAAGCAAAACACCTGGAGT 1007


```

VERSION CC548437.1 GI:31866721
KEYWORDS GSS.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS Holt,R., Cloutier,J., Yang,G., Barber,S., Smalilus,D., Prabhu,A.-L.,
Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., Keele,J.W. and Kappes,S.M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)
Other GSSs: CH240_432010.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Science Centre, Canada.
Plate: 432 row: O column: 10
Seq primer: 17
Class: BAC ends.

FEATURES
source
1. 836
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="bred: Hereford"
/db_xref="taxon:9913"
/clone="CH240_432010"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull LI Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match 22.3%; Score 798.2; DB 1; Length 836;
Best Local Similarity 97.2%; Pred. No. 0.16;
Matches 812; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 320 GCAACCACTTCAGTATCTTCCCTGTGAGACCCCATCAACAGCATGAAAGGCAAAATG 379
DB 2 GCAACCACTTCAGTATCTTCCCTGTGAGACCCCATCAACAGCATGAAAGGCAAAATG 61

QY 380 ATAGATATCTGAAGAGAACTCCCGAGTCAGTAGGTGGCCCATATGCTACTGAGATC 439
DB 62 ATAGATATCTGAAGAGAACTCCCGAGTCAGTAGGTGGCCCATATGCTACTGAGATC 121

QY 440 AGTGAGAAATTAATCTCCGAAGAATGAAGATGGAGCCCAAGCAAAAGAAATACCCAG 499
DB 122 AGTGAGAAATTAATCTCCGAAGAATGAAGATGGAGCCCAAGCAAAAGAAATACCCAG 181

QY 500 CTGTGGATGTGCTGGTGATATAAGCAAGGTCGGATGCTGTAAAGAGCAATATTGCAATG 559
DB 182 CTGTGGATGTGCTGGTGATATAAGCAAGGTCGGATGCTGTAAAGAGCAATATTGCAATG 241

QY 560 GAACCTGGAATGTCCAGGTCCATGAATCAAGGCAAAATTTGGAAGTGGTCAAAACAAGATGG 619
DB 242 GAACCTGGAATGTCCAGGTCCATGAATCAAGGCAAAATTTGGAAGTGGTCAAAACAAGATGG 301

QY 620 CAAGAGTGAATGTCACACATTCTAGGAATCAGCGAACTAAATGGACTGGATGGTGAAT 679

```

```

DB 302 CAAGAGTGAATGTCGACATTTAGGAATCAGCGAACTGAAATGGACTGGAATGGGTGAAT 361
QY 680 TTAACCTCAGATGACCATTTATATCTACTTCTGCGGCGAGGAATCCCTCAGAGAAATGGAG 739
DB 362 TTAACCTCAGATGACCATTTATATCTACTTCTGCGGCGAGGAATCCCTCAGAGAAATGGAG 421
QY 740 TAGCCATCATGTGTCAACAAAAGAGTCCGAAATGCAGTACTTGGATGCAATCTCAAAAACG 799
DB 422 TAGCCATCATGTGTCAACAAAAGAGTCCGAAATGCAGTACTTGGATGCAATCTCAAAAACG 481
QY 800 ACAGATGATCTCTGTTTCTTCCAGGCAAAACCATTCGAATATCAGATCAAGTAATCCAGTCT 859
DB 482 ACAGATGATCTCTGTTTCTTCCAGGCAAAACCATTCGAATATCAGATCAAGTAATCCAGTCT 541
QY 860 ATGCCCCCAACCAAGTATGCTGAAGAGCTGAAGTGAAGCTGCTATGAAGACCTTACAAG 919
DB 542 ATGCCCCCAACCAAGTATGCTGAAGAGCTGAAGTGAAGCTGCTATGAAGACCTTACAAG 601
QY 920 ACCTTTTAGAATCAACACCCAAAAGATGCTCTTCTCATTTAGGGGACTGGAATGCA 979
DB 602 ACCTTTTAGAATCAACACCCAAAAGATGCTCTTCTCATTTAGGGGACTGGAATGCA 661
QY 980 AAGTAGGAAGCAAGAAACACCTGGAGTAACAGGCAAAATTTGGCCCTTGGAAATACGGAATG 1039
DB 662 AAGTAGGAAGCAAGAAACACCTGGAGTAACAGGCAAAATTTGGCCCTTGGAAATACGGAATG 721
QY 1040 AAGCAGGGCAAGACATAATAGAGTTTTTGGCAAGAAATGCACTGGTCTATAGCAAAACCC 1099
DB 722 AAGCAGGGCAAGACATAATAGAGTTTTTGGCAAGAAATGCACTGGTCTATAGCAAAACCC 781
QY 1100 TCTTCCAAACACACAGAGAGACCTTACACATGACATCACCAGATGGTCAACA 1154
DB 782 TCTTCCAAACACACAGAGAGACCTTACACATGACATCACCAGATGGTCAACA 836

CC594700 938 bp DNA linear GSS 18-JUN-2003
CH240_396F13.77 CHORI-240 Bos taurus genomic clone CH240_396F13,
genomic survey sequence.
CC594700
CC594700.1 GI:31955684
GSS.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
Holt,R., Cloutier,J., Yang,G., Barber,S., Smalilus,D., Prabhu,A.-L.,
Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
Dalrymple,B.P. and Tellam,R.
Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other GSSs: CH240_396F13.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

```

British Columbia Genome Sciences Centre, Canada.
 Plate: 396 row: F column: 13
 Seq primer: T7
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..938
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_396F13"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

Query Match 22.3%; Score 798.2; DB 1; Length 938;
 Best Local Similarity 93.1%; Pred. No. 0.14; Mismatches 59; Indels 6; Gaps 4;
 Matches 876; Conservative 0

290 GACAGATGTGCTCCACTGGAGAGGAA-TGCAAAACCCTTCAGTATTTCTTGCTTGAG 348
 |||||
 938 GACACATGTGCTCCACTGGAGAGGAA-TGCAAAACCCTTCAGTATTTCTTGATTTAG 879
 |||||
 349 AACCCCATGAACAGTATGAAAAGCAAATGATAGGATATCTGAAAGAGAACTCCCCAGG 408
 |||||
 878 AA-CCCATGAACAGTATGAAAAGCAAATGATAGGATATGAAAGGAGAACTCCNCAGG 820
 |||||
 409 TCAGTAGTGCCCATATGCTACTGGAGATCAGTGGAGAATACTCCAGAAGAAATGAA 468
 |||||
 819 TCAGTAGTGCCCATATGCTACTGGAGATCAGTGGAGAATACTCCAGAAGAAATGAA 760
 |||||
 469 GAGATGGAGCCAAAGCAAAAAGATATCCAGCTCTGGATGTGACTGTGTATATAAGCAAG 528
 |||||
 759 GAGATGGAGCTAAAGCAAAAAGATATCCAGCTCTGGATGTGACTGTGTATATAAGCAAG 700
 |||||
 529 GTCCGATGCTTAAGAGCAATATTCATAGGAACCTTGGATGTGACTGTGTATATAAGCAAG 588
 |||||
 699 GTCCGATGCTTAAGAGCAATATTCATAGGAACCTTGGATGTGACTGTGTATATAAGCAAG 640
 |||||
 589 GGCAAAATGGAAATGGTCAAAACAGAGATGCAAGAGTGAATGTCAACATTTCTAGGAATC 648
 |||||
 639 GGCAAAATGGAAATGGTCAAAACAGAGATGCAAGAGTGAATGTCAACATTTCTAGGAATC 580
 |||||
 649 AGCGAACTAAATAGTGAATGGTGAATTTAACTCAGATGACCAATATATCTACTAC 708
 |||||
 579 AGCAAACTAAATAGTGAATGGTGAATTTAACTCAGATGACCAATATATCTACTAC 520
 |||||
 709 TCGGGCAGGATCCCTCAGAGAAATGGAGTACCATCATGTCAACAAAGAGTCCGA 768
 |||||
 519 TCGGGCAGGATCCCTTAGAAGAACTGGAGTAGCCATCATGTCAACAAAGAAATCAA 460
 |||||
 769 AATGAGTACTTGGATGTCAGTCTCAAAACGACAGAAATGATCTCTGTTGTTTCCAAAGGC 828
 |||||
 459 AATGAGTACTTGGATGTCAGTCTCAAAACGACAGAAATGATCTCTGTTGTTTCCAAAGGC 400
 |||||
 829 AAACCAATCAATATCAGATTAATCAAGTCTATGCCCCCAACAGTAATGCTGAAGAGCT 888
 |||||
 399 AAACCAATCAATATCAGATTAATCAAGTCTATGCCCCCAACAGTAATGCTGAAGAGCT 340
 |||||
 889 GAAATGGAACGGTCTCTATGAAGACCTCAAGACCTTTTATAGAACTAAACACCCCAAAAGAT 948
 |||||
 339 GAAATGGAACGGTCTCTATGAAGACCTCAAGACCTTTTATAGAACTAAACACCCCAAAAGAT 280
 |||||
 949 GTCCCTTCATATATAGGAGCTGGAATGCAAAAGTAGGAAGCAAGAAACACCTTGGAGTA 1008
 |||||
 279 GTCCCTTCATATATAGGAGCTGGAATGCAAAAGTAGGAAGTCA---AACACATGGAGTA 223
 |||||
 1009 ACAGGCAAAATTTGGCTTTGGAATCGGAATGAAGCAGGCAAGAACTAATAGAGTTTTCG 1068
 |||||
 222 ACAGGCAAAATTTGGCTTTGGAATCGGAATGAAGCAGGCAAGAACTAATAGAGTTTTCG 163
 |||||

QY 1069 CAAGAAATGCACTGGTTCATAGCAAAACCCCTCTTCCAAACAACAAGAGAAGCTCTAC 1128
 |||||
 DB 162 CAAGAAATGCACTGGTTCATAGCAAAACCCCTCTTCCAAACAACAAGAGAAGCTCTAC 103
 |||||
 QY 1129 ACATGACATCAACGATGGTCAACACCGAATCAGATTGATTATTTCTTGCAGCC-A 1187
 |||||
 DB 102 ACATGACATCAACGATGGTCAACACCGAATCAGATTGATTATTTCTTGCAGCCNA 43
 |||||
 QY 1188 AAGATGGAGAAAGCTCTATACAGTCAGCAAAAAACAAGACCAG 1228
 |||||
 DB 42 AAGATGGAGAAAGCTCTATACAGTCAGCAAAAAACAAGACCAG 2

RESULT 13

BZ882580
 LOCUS CH240_195M2_TV3 CHORI-240 Bos taurus genomic clone CH240_195M2,
 DEFINITION genomic survey sequence.

ACCESSION BZ882580

VERSION BZ882580.1 GI:29109992

KEYWORDS CSS.

SOURCE Bos taurus (cow)

ORGANISM

Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

1 (bases 1 to 922)

REFERENCE
 AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
 Shvartsbeyn,A., Gebregorgis,E., Chen,D., Riggs,F., de Jong,P.,
 Crawford,A.M. and McEwan,J.C.

TITLE Bovine BAC End Sequences from Library CHORI-240

JOURNAL

COMMENT

Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.html). This work

was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by AgResearch Ltd., New Zealand and The

Institute of Genomic Research (TIGR), USA.

Plate: 195 row: M column: 2

Seq primer: T7

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..922

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_195M2"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

Query Match 22.3%; Score 797.8; DB 1; Length 922;

Best Local Similarity 94.8%; Pred. No. 0.14; Mismatches 42; Indels 8; Gaps 5;

Matches 877; Conservative 0

QY 1692 ACATCTATTTCTCTTTATTTGACTATGCAAAAGCCCTTTGACTGGGGTCAACAATAAC 1751
 |||||

DB 1 ACATCTTTCTCTTTATTTGACTATGCAAAAGCCCTTTGACTGGGGTCAACAATAAC 60
 |||||

QY 1752 TGTGAAATTTCTGAAAGGATGGGAATACAGAGACCCTGACCTGACTCTTGGAAAAATT 1811
 DB 61 TGTGAAATTTCTGAAAGGATGGGAATACAGAGACCCTGACCTGACTCTTGGAAAAATT 120
 QY 1812 TGTATGTCAGGTGAGGAGCAACAGTGTAGAACTGGACATGGACAAACAGACTGGTTTCAAG 1871
 DB 121 TATATGTCAGGTGAGGAGCAACAGTGTAGAACTGGACATGGACAAACAGACTGGTTTCAAA 180
 QY 1872 TAGGAAAGGAGTGTCAAGGCTGTATATGTCACCGGCTGTGTAACTTCTATGAG 1931
 DB 181 TAGGAAAGGAGTGTCAAGGCTGTATATGTCACCGGCTGTGTAACTTCTATGAG 240
 QY 1932 AG-ACATCATGAGAAACCTGGCTGGAGGAGGACACAGCTGGAAATCAAGATTCGCGGA 1990
 DB 241 AGTACATCATGAGAAATGTGGCTGGAGGAGGACACAGCTGGAAATCAAGACTGCTGGGA 300
 QY 1991 GAAATAGCAATAACCTCAGATATGAGATGATACACCTTTATGGCAGAAAGTGAAGAGG 2050
 DB 301 GAAATATCAATAACCTCAGATACGAGATGATACACCTTTATGGCAGAAAGTGAAGAGG 360
 QY 2051 AACTTAAAGGCTCTTGATGAGGTCGAAAGGAGGAGTGAAGAGTGGCTTAAAGTCA 2110
 DB 361 AACTTAAAGGCTCTTGATGAGGTCGAAAGGAGGAGTGAAGAGTGGCTTAAAGTCA 420
 QY 2111 ACATTCAGAAACGAGATCATGTCATCTGCTCCCATCACTTCATGGAAATAGATGGG 2170
 DB 421 ACATTCAGAAACGAGATCATGTCATCTGCTCCCATCACTTCATGGAAATAGATGGG 480
 QY 2171 AACAGTGGAAACAGTGTGAGCTTATTTTGGGGGGCTCCAAATCACTGCAGATGG 2230
 DB 481 AACAGTGGAAACGAGTGTGAGCTTATTTTGGGGGGCTCCAAATCACTGCAGATGG 538
 QY 2231 TGACTGCGACCATGAAATTTAAAGACCTTACTCTCTGGAAGAAAGTTA--ACCAACCT 2288
 DB 539 TGACTGCGACCATGAAATTTAGAGACGCTTACTCTCTGGAAGAAAGTTATGACTTAACCT 598
 QY 2289 AGATAGCATATTGAAAGCAGAGACATTAATCTTCCCAACAAAGCCCATCTAGTCAAGGC 2348
 DB 599 AGATAGCATATTGAAAGCAGAGACATTAATCTTCCCAACAAAGCCCATCTAGTCAAGGC 658
 QY 2349 TATGTTTTTCCAGTGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2408
 DB 659 TATGTTTTTCCAGTGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 718
 QY 2409 CTGAGAAATGATGCTTTTGAACCTGTGCTGTGAGAGAGACTCTTGAAGAGTCCCTTGGAC 2468
 DB 719 CCGAAGAAATGATGCTTTTGAACCTGTGCTGTGAGAGAGACTCTTGAAGAGTCCCTTGGAC 778
 QY 2469 TGCAAGGAGATCCACAGTCCATCTCAAGAGATCAGCCCTGGGATTTCTTGGAGG 2528
 DB 779 TGCAAGGAGATCCACAGTCCATCTTAAGAGATCAGCCCTGGGATTTCTTGGAGG 838
 QY 2529 AATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGAAGAGTGAATCACT 2588
 DB 839 AATGATGCTAAAGCTGAAACT--CAGTACTTTGGCCACCTCATGCGAAGAGTGAATCACT 897
 QY 2589 GGAAAGACCTGATGCTGGAGGGAT 2615
 DB 898 GGAAAGA--TTGATGCTGGAGGGAT 922

RESULT 14
 CC480115/c 899 bp DNA linear GSS 16-JUN-2003
 LOCUS CH240_30704.77 CHORI-240 Bos taurus genomic clone CH240_30704,
 DEFINITION genomic survey sequence.
 ACCESSION CC480115
 VERSION CC480115.1 GI:31759328
 KEYWORDS GSS.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

REFERENCE
AUTHORS

Bovidae; Bovinae; Bos.
 1 (bases 1 to 899)
 Holt,R., Scott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
 Tsai,M., Cloutier,A., Lee,D., Girm,N., Olson,T., Mayo,M.,
 Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
 Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
 Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
 Dalrymple,B.P. and Tellam,R.
 Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 Unpublished (2003)
 Other GSSs: CH240_30704.TARBAC13P2
 Contact: Rob Holt
 Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bcgsc.ca

TITLE
JOURNAL
COMMENT

Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 307 row: O column: 4
 Seq primer: T7

Class: BAC ends.
 Location/Qualifiers

FEATURES
source

1..899
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_30704"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull 11 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 22.3%; Score 797.2; DB 1; Length 899;
 Best Local Similarity 94.5%; Pred. No. 0.15;
 Matches 826; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 725 TCAGAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCGAATGCAGTACTGGAT 784
 DB 899 TCAGAGAAATGGAGTAGCCATCATGGTCAACAAAGAGTCCGAATGCAGTACTGGAT 840
 QY 785 CGAGTCTCAAAACGACAGATGATCTCTGTTGTTTCCAGGGCAACCATCAATATCA 844
 DB 839 GCAATCTCAAAATGACAGCAATATCTGTTTCGTTTCCAGGGCAACCATCAATATCA 780
 QY 845 CAGTATCAAGTCTATGCCCAACAGTAACTGCTGAAGAGTCAAGTGAACGCTCT 904
 DB 779 CAGTATCAAGTCTATGCCCAACAGTAACTGCTGAAGAGTCAAGTGAACGCTCT 720
 QY 905 ATGAAGACCTCAAGACCTTTTAGAACTAAACCCCAAAAAGATGCTCTCTCATATAG 964
 DB 719 GTGAAGAACTCAAGACCTTTTAGAACTAAACCCCAAAAAGTGTCTCTCATATAG 660
 QY 965 GGGACTGGAATCGAAAGTAGGAAGCAAGAAACACCTGGAGTAAACGGCAAAATTGGCC 1024
 DB 659 GGGACTGGAATCGAAAGTAGGAAGTCAAGAAACACCTGGAGTAAACGGCAAAATTGGCC 600
 QY 1025 TTGGAATACGGAATGAAGCAGGGGCAAGACTAATAGAGTTTTCGCAAGAAATGCACCTG 1084
 DB 599 TTGGAATACGGAATGAAGCAGGGGCAAGACTAATAGAGTTTTCGCAAGAAATGCACCTG 540
 QY 1085 TCATAGCAAAACCTCTTCCCAACCAACAGAGAAAGCTCTACACATGGACATCACCAG 1144

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada.

Plate: 415 row: B column: 11

Seq primer: SP6

Class: BAC ends.

FEATURES

Location/Qualifiers

1..904
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="bred: Hereford"
/db_xref="taxon:9913"
/clone="CH240_415B11"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

Query Match 22.1%; Score 790.8; DB 1; Length 904;

Best Local Similarity 94.7%; Pred. No. 0.15;

Matches 859; Conservative 0; Mismatches 42; Indels 6; Gaps 4;

QY	289	TCAGAGATGTGTCCTAGTGAAG--GGGATCGAAACACCTTCAGTATCTTCGCTTG	346
DB	1	TCAGAGATGTGTCCTAGTGAAGCGTAAATGGCAACACCTTCAGTATCTTCGCTTG	60
QY	347	AGAACCCCATGACAGTATGAAGGCAAAATGATAGTACTGAAAGAGAACTCCCCA	406
DB	61	AGAACCCCTTGAACAGTATGAAGGCAAAATGATAGTACTGAAAGAGAACTCTCA	120
QY	407	GTCAGTAGTGTGTCCTAGTGAAGTCACTGAGATCACTGAGGAAATTAATCCAGAAAGATG	466
DB	121	GATCAGTAAGGGCCCAATAGTCTACTGAGATCGGTGAGAAATTAATCCAGAAAGATG	180
QY	467	AAGAGATGGAGCCCAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGGTGATATAGCA	526
DB	181	AGGGATGGAGCCCAAGCAAAAGAAATACCACTGTGGATGTGACTGGTGATATAGCA	240
QY	527	AGTCCGATGTGTAAGAGCAATATTCATAGGAACCTCGAATGTGAGTCCATGAATC	586
DB	241	AGTCCGATGTGTAAGAGCAATATTCATAGGAACCTCGAATGTGAGTCCATGAATC	300
QY	587	AAGCAAAATGGAGTGTGCAACAGAGATGCAAGATGCAATGTCAACATCTTAGGA	646
DB	301	AAGCAAAATGGAGTGTGCAACAGAGATGCAAGATGCAATGTGCAATCTTAGGA	360
QY	647	TCAGCAACTAAATGGACTGGAATGGGTGAATTTAACTCAGATGACCAATATATCTACT	706
DB	361	TCAGCAACTAAATGGACTGGAATGGGTGAATTTAACTCAGATGACCAATATATCTACT	420
QY	707	ACTCGGGCAGGATCCCTCAGAGAAATGGAGTGGATGATGTCATGTCACAAAGATCC	766
DB	421	ACTCGGGCAGGATCCCTCAGAGAAATGGAGTGGATGATGTCATGTCACAAAGATCC	478
QY	767	GAATGCGAGTACTGGATGATGATCTCAAAACACAGATGATCTCTGTTGTTTCCAAAG	826
DB	479	GAATGCGAGTACTGGATGATGATCTCAAAACACAGATGATCTCTCTCTGTTTCCAAAG	538
QY	827	GCAACCAATTCATATCAGATGAATCAAGTCTATGCCCCCAACAGTATGCTGGAAGAG	886
DB	539	GCAACCAATTCAGTATCAGATGAATCAAGTCTATGCCCCCAACAGTATGCTGGAAGAG	598
QY	887	CTGAAGTTGAACCGTCTCTATGAAGACTCAACAGCTTTTAGAAGTAAACACCCCAAAAG	946
DB	599	CTGAAGTTGAACCGTCTCTATGAAGACTCAACAGCTTTTAGAAGTAAACACCCCAAAAG	658
QY	947	ATGTCCTTCTATATAGGGAGTGGAAATGCAAAAGTAGGAAGCAAGAAACACCTGGAG	1006

DB	659	ATGTCCTTTTCATTATATAGGGAGTGGAGTCAAAAGTAGGAAGTCAAGAAACACCTGGAG	718
QY	1007	TAAACAGCAAAATTTGGCTTGGAAATACGGAATGAAGCAGGCAAGAGACTAAT-AGAGTTT	1065
DB	719	TAAACAGCAAAATTTGGCTTGGAAATACGGAATGAAGCAGGCAAGAGACTAATAGAGTTT	778
QY	1066	TGCCAAGAAATATGCACTGTGTATAGCAAAACACCTCTTCCAAACAACAAGAGAGACTC	1125
DB	779	TGCCAAGAAATATGCACTGTGTATAGCAAAACACCTCTTCCAAACAACAAGAGAGACTC	838
QY	1126	TACACATGACATCACCATGCTCAACACCAAAATCAGATGATTATATCTTTTGCAGC	1185
DB	839	TATACATGACATCACCATGCTCAACACCAAAATCAGATGATTATATCTTTTGCAGC	897
QY	1186	CAAAGAT 1192	
DB	898	CAAAGAT 904	

RESULT 19

CC476492

LOCUS

DEFINITION

CH240_30208.T7 CHORI-240 Bos taurus genomic clone CH240_30208,

genomic survey sequence.

ACCESSION

CC476492

VERSION

CC476492.1

GI:31753755

KEYWORDS

GSS.

SOURCE

Bos taurus (cow)

ORGANISM

Bos taurus

REFERENCE

1 (bases 1 to 897)

Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A., L.,

Tsai, M., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M.,

Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,

Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,

Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,

Dalrymple, B. P. and Tellam, R.

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

Unpublished (2003)

Other_GSSs: CH240_30208.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave. Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bccsc.ca

Clones are derived from the bovine BAC library CHORI-240

(<http://www.chori.org/bacpac/bovine240.htm>). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/ordering_information.htm). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 302 row: 0 column: 8

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..897

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="bred: Hereford"

/db_xref="taxon:9913"

/clone="CH240_30208"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

FEATURES

source

Query Match		22.1%;	Score 790.2;	DB 1;	Length 897;
Best Local Similarity		95.1%;	Pred. No. 0.16;		
Matches 856;		Conservative	0;	Mismatches	38; Indels 6; Gaps 4;
Qy	1664	ATGGAAAAAGCAAGAGAGTTCAGAAAAACATCTATTTCGCTTTATTGACTATGCACAAA	1723		
Db	1	ATGGAAAAAGCAAGAGAGTTCAGAAAAACCATCTATTTCGCTTTATTGACTATGCACAAA	60		
Qy	1724	GCCTTTGACTGTGGGGGTGCACATAAACTGTGAAAAATTTCTGAAAGGGATGGGAATACCA	1783		
Db	61	GCCTTTGACTGTGGATGCACATCAACTGTGAAAAATTTCTGAAAGAGATAGGAATACCA	120		
Qy	1784	GACCACCTGACCTGACTCTTTGAAAAATTTGATGCGAGTCAGGACAGCAACAGTTAGAACT	1843		
Db	121	GACCACCTGATCTGCTCTTTGAGAAATCTGTATGCGAGTCAGGACAGCAACAGTTAGAACT	180		
Qy	1844	GGACATGGAAACACAGACTGGTTCCAAGTAGGAAAAAGGAGTAGTGTCAAGGCTGTATATTG	1903		
Db	181	GGACATGGAAACACAGACTGGTTCCAAATAGGAAAAAGGACTTCGTCAAGGCTGTATATTG	240		
Qy	1904	TCACCCGGCTGTTTAACTTCTATGCAGAG-ACATCATGAGAAAAAGCTGGGCTGGAAGAA	1962		
Db	241	TCACCCGTCTTATTTAATTTATATGCAGAGTACATCATGAGAAAAAGCTGGGCTGGAAGAA	300		
Qy	1963	GCACAGCTGGAATCAAGATTGCGGGAGAAATAGCAATAACCTCAGATATGCAGATGAT	2022		
Db	301	GCACAGCTTGAATCAAGATTGCGGGAGAGAAATATCAATAACCTCAGATATGCAGATGAT	360		
Qy	2023	ACCACCTTTATGCGCAAAAGTGAAGAGAACTAAAAAGCCTTTGATGAAGTGAAGAG	2082		
Db	361	ACCACCTTTATGCGCAAAAGTGAAGAGAACTAAAAAGCCTTTGATGAAGTGAAGAG	420		
Qy	2083	GAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAACGAGATCATGGCATCTGT	2142		
Db	421	GAGAGTGAAGAAAGTTGGCTTAAAGCTCAACATTCAGAAAAACGAGATCATGGCATCTGT	480		
Qy	2143	CCCATCACTTCATGGGAAATAGATGGGGAAAACAGTGGAAACAGTGTGACACTTTATTTT	2202		
Db	481	CCCATCACTTCATGGGAAATAGATGGGGAAAACAGTGGAAACAGTGTGACACTTTATTTT	540		
Qy	2203	TGGGGGGCTCCAAAATCACTGCAGATGGTGACTGCGAGCCATGAAATTAAGACACTTAC	2262		
Db	541	--CTGGGGCTCCAAAACCACTGCAGATGGTGACTGTGTAGCCATGAAATTAAGACACTTAC	598		
Qy	2263	TCCTTGGAGAAAAAGTTA--ACCAACTAGATAGCATATTGAAAAACAGAGACATTACCT	2320		
Db	599	TCCTTGGAGAAAAAGTTATGACCACTAGATAGCATATTCAAAAGCAGAGACATTACTT	658		
Qy	2321	TGCCAACAAGCCCCATCTAGTCAAGGCTATGGTTTTTCAGTGGCTCATGTATGGATGTG	2380		
Db	659	TGCCAACAAGAGTTGCTGTAGTCAAGGCTATGGTTTTTCCTGTGGTCAATGTATGGATGTG	718		
Qy	2381	AGAGTTGGACTGTGAAGAAAGCTGAGCACTGAAAGAAATTGATGCTTTTGAACCTGTGTGT	2440		
Db	719	AGAGTTGGACTGTGAAGAAAGCTGAGCACTGAAAGAAATTGATGCTTTTGAACCTGTGTGT	778		
Qy	2441	GGAGAGACTCTTTGAGAGTCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCGAAAGG	2500		
Db	779	GGAGAGACTCTTTGAGAATCCCTTGGACTGCAAGGAGATCCAAACAGTCCATTCGAAAGG	838		
Qy	2501	AGATCAGCCCTGGGATTTCTTTGGAGGAATGATGCTAAAGCTGAAACCTCAGTACTTTG	2560		
Db	839	AGATCAGTCTCGGATTTCTTTGGAA--GAATGATGCTAAAGCTGAAACCTCAGTACTTTG	897		
RESULT 20					
CC579657/c					
LOCUS					
DEFINITION					
CH240_374G1.T7 CHORI-240 Bos taurus genomic clone CH240_374G1,					
genomic survey sequence.					
ACCESSION					
CC579657.1 GI:31925542					
VERSION					

KEYWORDS	GSS.
SOURCE	Bos taurus (cow)
ORGANISM	Bos taurus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
AUTHORS	1 (bases 1 to 881) Holt, R., Stott, J., Yang, G., Barber, S., Smalilus, D., Prabhu, A.-L., Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M., Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S., Schein, J., Marra, M., de Jong, P., McWilliam, S., Barrie, W., Dalrymple, B. P. and Tellam, R.
TITLE	Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
JOURNAL	Unpublished (2003)
COMMENT	Other_GSSs: CH240_374G1.TARBA13P2 Contact: Rob Holt Sequencing The British Columbia Cancer Agency Genome Science Centre 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6 Tel: 604-877-6085 Fax: 604-877-6276 Email: rholt@bcgsc.ca Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada. Plate: 374 row: G column: 1 Seq primer: T7 Class: BAC ends.
FEATURES	Location/Qualifiers 1.. 881 /organism="Bos taurus" /mol_type="genomic DNA" /strain="breed: Hereford" /db_xref="taxon:9913" /clone="CH240_374G1" /sex="Male" /cell_type="Blood" /clone_lib="CHORI-240" /note="Vector: PTARBAC1.3; Site 1: Mbol; Site 2: Mbol; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"
Query Match	22.0%; Score 785.2; DB 1; Length 881;
Best Local Similarity	95.0%; Pred. No. 0.17;
Matches 811; Conservative	0; Mismatches 43; Indels 0; Gaps 0;
Qy	745 ATCATGTCACAAAGAGTCCGAAATGCAGTACTTGGATGCAGTCTCAAAACGACAGA 804
Db	881 ATCATGTCACAAAGAGTCCGAAATGCAGTACTTGGATGCAGTCTCAAAACGACAGA 822
Qy	805 ATGATCTCTGTTTTCAGAGCAACCAATTCATATCACAGTATCCAGTCTATGCC 864
Db	821 ATGATCTCTGTTTTCAGAGCAACCAATTCATATCACAGTATCCAGTCTATGCC 762
Qy	865 CCAACCAAGTAATGCTCAAGAGCTGAGTTGAACGGTCTATCAAGACCTTACAGACCTT 924
Db	761 CCAACCAAGTAATGCTCAAGAGCTGAGTTGAACGGTCTATCAAGACCTTACAGACCTT 702
Qy	925 TTAGAACTAACCCCAAAAAAGATGTCCTTCTCATTTATAGGGGACTGGAATGCAAAAGTA 984
Db	701 TTAGAACTAACCCCAAAAAAGATGTCCTTCTCATTTATAGGGGACTGGAATGCAAAAGTA 642
Qy	985 GGAAGCAAAAGAAACACCTGGAGTAAACAGGCAATTTGGCTTTGGAAATACGGAATGAAGCA 1044
Db	641 GGAAGTCAAGAAACACCTGGAGTAAACAGGCAATTTGGCTTTGGAAATACAGAAATGAAGCA 582
Qy	1045 GGGCAAGAGACTAATAGAGTTTTCGCAAGAAATGCAGTGTGTCATAGCAAAACACCCCTCTTC 1104

Db 581 GGGCAAGACTAATAGAGTTTGGCAAGAAATGACCTGATCATAAACAACACCTCTTC 522
 QY 1105 CAACAACAAGAGAGAGACTCTACACATGGACATCACAGATGCTCAACACCGAAATCAG 1164
 Db 521 CAACAACAAGAGAGAGACTCTATACATGGATATCAACAGATGCTCAACACCGAAATCAG 462
 QY 1165 ATTGATTATATCTTTGAGCAAGATGGAGAGCTCTATACAGTACGAAACAAAGA 1224
 Db 461 ATTGATTATATCTTTGAGCAAGATGGAGAGCTCTATACAGTACGAAACAAAGA 402
 QY 1225 CCAGGAGCTTACTGTGGCTCAGATCATGAATCTCTTATGGCAATTCAGACTTAAATG 1284
 Db 401 CCAGGAGCTTACTGTGGCTCAGACCATGAATCTCTTATGGCAATTCAGACTTAAATG 342
 QY 1285 AAGAAAGTAGGAAACCACTAGATCATCTAGGTAAGACCTTAATCCAAATCCCTTATGAT 1344
 Db 341 AAGAAAGTCGGAAACCACTAGACCATTCAGGTATGACCTTAATCCAAATCCCTTATGAT 282
 QY 1345 TATACAGTGAAGTGAAGAAATAGATTTAAGGGCTTAGATCTGTAGTACAGAGTACCTAAT 1404
 Db 281 TATACAGTGAAGTGAAGAAATAGATTTAAGGGCTTAGATCTGTAGTACAGAGTACCTAAT 222
 QY 1405 GAACATATGACAGAGTTTCATGACATTTGACATGTTACAGGAGAGAGGATCGAGACCATCCCATG 1464
 Db 221 GAACATATGACAGAGTTTCATGACATTTGACATGTTACAGGAGAGAGGATCGAGACCATCCCATG 162
 QY 1465 GAAAGAAATCAAAAAGCAAAATGCTGTCTGGGAGGCTTACAAATAGCTGTGAAA 1524
 Db 161 GAAAGAAATCAAAAAGCAAAATGCTGTCTGGGAGGCTTACAAATAGCTGTGAAA 102
 QY 1525 AGAAGAGAGTCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGTATCAATCGAGAGTTC 1584
 Db 101 AGAAGGAGTCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGTATCAATCGAGAGTTC 42
 QY 1585 CAAGAAGTTCAG 1598
 Db 41 CAAGAATAGCAAG 28

 RESULT 21
 CC500971
 LOCUS CH240_339C15.25 bp DNA linear GSS 17-JUN-2003
 DEFINITION CH240_339C15.25 CHORI-240 Bos taurus genomic clone
 CC500971
 ACCESSION CC500971
 VERSION GSS
 KEYWORDS CC500971.1 GI:31819264
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoides;
 Bovidae; Bovinae; Bos.
 1 (bases 1 to 925)
 Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
 Tsai, M., Cloutier, A., Lee, D., Gilm, N., Olson, R., Mayo, M.,
 Butterfield, J., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B. P. and Tellam, R.
 Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 Unpublished (2003)
 TITLE Other_GSSs: CH240_339C15.77
 JOURNAL Contact: Rob Holt
 COMMENT Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bccsc.ca
 Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdjong@mail.chori.org).

Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 339 row: C column: 15
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 1..925
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_339C15"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBA1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull, Li Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 22.0%; Score 784.6; DB 1; Length 925;
 Best Local Similarity 93.0%; Pred. No. 0.16;
 Matches 863; Conservative 0; Mismatches 59; Indels 6; Gaps 4;

QY 1796 TGACTCTTCAAAAATTTGTATGACGTCAGGAGCAACAGTTAGAACTGGACATGGAACA 1855
 Db 1 TGCCTCTTGAAGAAATTTGTATGACCTCATGAAGCAACAGTTAGAACTGGACATGGAACA 60
 QY 1856 ACAGACTGCTTCCAAAGTAGGAAAGAGTATGTCGAAGCTGTATATTCTACCCGGCTTG 1915
 Db 61 ACAGACCCTTCCAAATAGGAAAGAGTTCGTCAAGCTGTATATTCTACCCCTGTTA 120
 QY 1916 TTTAACTCTTATGACAG-ACATCATGAGAAACGCTGGCTGGAGAAAGCAAGCTGGA 1974
 Db 121 TTTAACTCTTATGACAGTACATCATGAGAACTCTGGACTGGAGAAACCAAGCTGGA 180
 QY 1975 ATCAAGATTGCGGAGAAATAGCAATTAACCTCAGATATGAGATGATGACACCCCTTATG 2034
 Db 181 ATCAAGATTGCTGGAGAAATATCAATTAACCTCAGATATGAGATGATGACACCCCTTATG 240
 QY 2035 GCAGAAAGTAGGAGAACTAAAGCCCTTGTATGAAGGTGAAAGAGGAGAGTGAAGAA 2094
 Db 241 GCAGAAAGTAGGAGAACTAAAGAGCTTGTATGAAGTGAAGTGAAGTGAAGTGAAGAA 300
 QY 2095 GTTGCTTAAAGCTCAACATTCAGAAACAGAGATCATGGCATCTGGTCCCATCACTTCA 2154
 Db 301 GTTGCTTAAAGCTCAACATTCAGAAACAGAGATCATGGCATCTGGTCCCATCACTTCA 360
 QY 2155 TGGGAAATAGATGGGAAACAGTGGAAACAGTGTGACAGACTTTATTTTGGGGGGCTTCA 2214
 Db 361 TGTAAATAGATGGGAAACAGTGGAAACAGTGTGACAGACTTTATTTT-CTGGGCTCA 418
 QY 2215 AAATCAGTGCAGATGCTGACATGACCCATGAATTAAGACACTTACTCTCTGGAAGAA 2274
 Db 419 AAATCAGTGCAGATGCTGACATGACCCATGAATTAAGAGATGTTTACTCTCTGGAAGAA 478
 QY 2275 AAGTTA--ACCAACCTAGATGACATATTGAAAGCAGAGACATTAACCTTGGCAACAAGC 2332
 Db 479 AAGTTATACCAACCTAGATGACATATTGAAAGCAGAGACATTAACCTTGGCAACAAGC 538
 QY 2333 CCCATCTAGTCAAGCTATGTTTTTCCAGTGGTCTGATGATGATGATGATGATGATGATG 2392
 Db 539 TTTGCTAGTCAAGCTATGTTTTTCCAGTGGTCTGATGATGATGATGATGATGATGATG 598
 QY 2393 TGAAGAAAGCTGAGCAGCTGAAGAAATGATGCTTTTGAACCTGCTGCTGCTGCTGCTGCT 2452
 Db 599 TGAAGAGGCTGAGGCGGGAAGAAATGATGCTTTTGAACCTGCTGCTGCTGCTGCTGCT 659
 QY 2453 TGAGAGTCCCTTGGAGTGCAGAGGAGATCCCAACCAAGTCCATTCTGAAGGAGATCAGCCCTG 2512
 Db 659 TGAGAGTCCCTTGGAGTGCAGAGGAGATCCCAACCAAGTCCATTCTGAAGGAGATCAGCCCTG 718

```

QY 2513 GGATTTCTTTGGAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGTATCA 2572
Db 719 GGATTTCTTTGGAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGTATCA 778
QY 2573 GAAGAGCTGACTCAGTGGGAAAGACCTCTGATGCTGGAGGGATTTGGGGGACAGGAGAGAA 2632
Db 779 GAAGAGTTGACCCATTGGAAAGACCTCTGATGCTGGAGGGATTTGGGGGACAGGAGAGAA 838
QY 2633 GGGGACAGCAGAGGATGAGATGCTGGATGGCATCACTGACTCGATGGAGCGTGAATCTGG 2692
Db 839 -GGGACAATAGAGGATGAGATGCTGGATGGCATCACTGACTCGATGGAGCGTGAATCTGG 897
QY 2693 GTCAACTCTCTGGAGTTGGTGATGGACAG 2720
Db 898 GTGAACTCCGGAGTTGGTGATGGACAG 925

RESULT 22
CC543667/c
LOCUS
DEFINITION
CH240_425B21..T7 CHORI-240 Bos taurus genomic clone CH240_425B21,
genomic survey sequence.
ACCESSION
CC543667
KEYWORDS
GSS.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 934)
Holt,R., Stott,J., Yang,G., Barber,S., Smalrus,D., Prabhu,A.-L.,
Teal,M., Cloutier,A., Lee,D., Girm.N., Olson,T., Mayo,M.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
Mathewson,C., Wye,N., Mason,A., Brown-John,M., Jones,S.,
Schein,J., Marra,M., de Jong,P., Keele,J.W. and Kappes,S.M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)
Other GSSs: CH240_425B21.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bccsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 425 Row: B Column: 21
Seq primer: T7
Class: BAC ends.
FEATURES
Location/Qualifiers
1..934
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_425B21"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match 22.0%; Score 784.2; DB 1; Length 934;

```

```

Best Local Similarity 93.0%; Pred. No. 0.16;
Matches 873; Conservative 0; Mismatches 58; Indels 8; Gaps 5;
QY 1812 TGTATGAGGCTCAGGAAAGCAACAGTGTAGAACTGGACATGGAACAACAGACTGGTTTCAAG 1871
Db 934 TGTATGAGGCTCAGGAAAGCAACAGTGTAGAACTGGACATGGAACAACAGACT-GTTTCAA 876
QY 1872 TAGGAAAGAGAGTATGTCAAAGCTGTATATTTGTCAACGGCTGTTTAACTTCTATGCGAG 1931
Db 875 TAGGAAAGAGAGTATGTCAAAGCTGTATATTTGTCAACGGCTGTTTAACTTCTATGCGAG 816
QY 1932 AG-ACATCATCAGAAACCGCTGGGCTGGAAGAAGCAACAGCTGGGAATCAAGATTCCCGGGA 1990
Db 815 AGTACATTATGAGAAACCGCTGGACTGGAGAAACACAGCTGGGAATCAAGATTCCCGGGA 756
QY 1991 GAAATAGCAATAACCTCAGATATATGCAAGATATACACCCCTTATGGCAGAAAGTGAAGAG 2050
Db 755 GAAATATCAATAACCTCAGATATATGCAAGAGACACACCCCTTATGGCAGAAAGTGAAGAG 696
QY 2051 AACTAAAAAGCCTCTTGATGAGGTGAAGAGGAGAGTGAAAAGTTTGGCTTAAAGCTCA 2110
Db 695 AACTAAAAAGCCTCTTGATGAGGTGAAGAGGAGAGTGAAAAGTTTGGCTTAAAGCTCA 636
QY 2111 ACATTTCAGAAAACGAGATCATGGCATCTGGTCCCATCACTTCATTCATGGGAAATAGATGGG 2170
Db 635 ACATTTCAGAAAACGAGATCATGGCATCTGGTCCCATCACTTCATTCATGGGAAATAGATGGG 576
QY 2171 AAACAGTGAAACAGTGTCAAGCTTTATTTTGGGGGGCTCCAAAATCACTGCGAGATGG 2230
Db 575 AAACAGTGAAACAGTGTCAAGCTTTATTTTCTGGGCTCCAAAATCACTTACAGATGG 518
QY 2231 TGACTGCGACCATGAAATTTAAAGACACTTACTCTCTGGAGAAAAGTTA--ACCAACT 2288
Db 517 TGATTGCGAAGTAAGAAATTTAAAGATGCTTACTCTCTGGAGAAAAGTTATGACCACT 458
QY 2289 AGATAGCATATTGAAAAGCAGACAGATTAATCTTTGCCAACAAAGCCCTCTAGTCAAGGC 2348
Db 457 AGATAGCATATTGAAAAGTAGACATTAATCTTTGCCAACAAAGGTTTCTAGTCAAGGC 398
QY 2349 TATGTTTTTCCAGTGGTCTATGATGATGAGAGTGGACTGTGAAGAAAGCTGAGCA 2408
Db 397 TATGTTTTTCTGTTGTCATGATGATGATGAGAGTGGACTGTGAAGAAAGCTGAGCA 338
QY 2409 CTGAAAGATTGATGCTTTTGAACCTGTGGTGTGGAGAAAGACTCTTTGAGAGTCCCTTGGAC 2468
Db 337 CCAGAGATTGATGCTCTTGAACCGCGTGTGGAGA--ACTCTTGAGAGTCCCTTGGAC 280
QY 2469 TGAAGGAGATCCAAACAGTCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGG 2528
Db 279 TGAAGGAGATCCAAACAGTCCATTTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGG 220
QY 2529 AATGATGCTAAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCAGAAAGCTGACTCACT 2588
Db 219 AATGATGCTAAAAGCTGAAACTCCAGTACTTTGGCCACCTCATCGAAGAGTTGACTCACT 160
QY 2589 GGAAGAAAGCCCTGATCTCGGAGGGATTTGGGGCAGGAGAGAGAAAGGGGACGACAGAGAT 2648
Db 159 GGAAGAAAGCTCTGATCTGAGAGGGATTTGGGGCAGGAGAGAGGGGATGACAGAGAT 100
QY 2649 GAGATGCTGGATGGCATCACTGATGAGAGTGGAGTCTGGGTGAACCTCTCTGAGTT 2708
Db 99 GAGATGCTGGAGGGCATCACTGATGAGATGAGATGAGTCTGAGTGAATCCAGAGAT 40
QY 2709 GGTGATGGACAGGAGGCGCTGTCTCGCGCGGATTCATGG 2747
Db 39 GGTGATGGACAGGAGGCGCTGTCTGAGATTCATGG 1

```

```

RESULT 23
CC547247
LOCUS
DEFINITION
CH240_430B11.TARBAC13P2 CHORI-240 Bos taurus genomic clone
CH240_430B11, genomic survey sequence.

```

CC547247
CC547247.1 GI:31865531
GSS
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 907)
Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., Keele, J.W. and Kappes, S.M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)
Other GSSs: CH240_430B11.77
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6065
Fax: 604-877-6276
Email: rholtebcsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Science Centre, Canada.
Plate: 430 row: B column: 11
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..907
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_430B11"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="vector: pFARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match 21.9%; Score 783.8; DB 1; Length 907;
Best Local Similarity 94.1%; Pred. No. 0.16;
Matches 85; Conservative 0; Mismatches 48; Indels 6; Gaps 4;
library (Male) produced by Pieter de Jong"

1796 TGACTCTGAAAAATTGTATGTCAGGTGTCAGGAGCAACAGTTAGAACTGGACATGGAACA 1855
1 TGCTCTTGAGAAATTTGTATGTCAGGTGTCAGGAGCAACAGTTAGAACTGGACATGGAACA 60
1856 ACAGACTGTTCCAGTATGAGAAAGGATGTATGTCAGGCTGTATATGTACCCGCTTG 1915
61 ACAGACTGTTCCAGTATGAGAAAGGATGTATGTCAGGCTGTATATGTACCCCTGTTTA 120
1916 TTTAACTTCTATGCAGAG-ACATCATGAGAAAGCTGGGCTGGAAGAACACAAAGCTGGA 1974
121 TTTAACTTATATGCAGAGTACATCATGAGAAAGC-GGACTGGAAGAAACACAAAGCTGGA 179
1975 ATCAAGATTCCCGGAGAAATACATTAACCTCAGATATGCAGATGATACCACTTATG 2034
180 ATCAAGATTCCCGGAGAAATATCAATCACTCAGATATGCAGATGATACCACTTATG 239
2035 GCAGAAAGTGAAGAGGAACCTTAAAGCCCTCTTATCAAGGTGAAGAGGAGAGTGAAGAA 2094
240 GCAGAAAGTGAAGAGGAACCTTAAAGCCCTCTTATCAAGGTGAAGAGTGAAGAGTGAAGAA 299

2095 GTTGGCTTAAAGCTCAACATTCAGAAAACGAAAGATCATGGCATCTGGTCCCATCACTTCA 2154
300 GTTGGCTTAAAGCTCAACATTCAGAAAACGAAAGATCATGGCATCTGGTCCCATCACTTCA 359
2155 TGGGAATAGATGGGAAAACAGTGGAAAACAGTGTGAGACTTTTATTTTGGGGGGCTCCA 2214
360 TGGGAATAGATGGGAAAACAGTGGAAAACAGTGTGAGACTTTTATTTT--CTGGGCTCCA 417
2215 AATCATCTCAGATGGTGAAGTGGAGCCATGAAATTAAGACACACTTACTCTTGGAGAA 2274
418 AATCATCTCAGATGGTGAAGTGGAGCCATGAAATTAAGACACACTTACTCTTGGAGAA 477
2275 AGTTTA--ACCAACCTTAGATGATGAAAGAGAGACACATTACCTTGGCAACAAAGC 2332
478 AGTTATGACCAACCTTAGATGATGATGAAAGAGAGACACATTACTTTGCCAACAAAGG 537
2333 CCCATCTAGTCAAGGCTATGGTTTTTCCAGTGGTGTATGATGATGATGATGATGATGATG 2392
538 TTCTGCTAGTCAAGGCTATGGTTTTTCTGTGCTCATGTATGATGATGATGATGATGATG 597
2393 TGAAGAAAGCTCAGCACTGAGAAATGATGCTTTTGAAGTGTGTTGGAGAGACTCT 2452
598 TGAAGAAAGCTCAGCGCCAAATAATGATGCTTTTGAAGTGTGTTGGAGAGACTCT 657
2453 TGAGAGTCCCTTGGAGTCAAGGAGATCCAAACAGTCCATTTCTGAAGAGATCAGCCCTG 2512
658 TGAGAGTCCCTTGGAGTCAAGGAGATCCAAACAGTCCATTTCTGAAGAGATCAGCCCTG 717
2513 GAATTTCTTGAAGAGATGATGCTTAAAGTGAAGTCAAGTCCAGTACTTTGGCCACTGATCA 2572
718 GGATATCTTTGGAAGAGTGTGCTTAAAGTGAAGTCAAGTCCAGTACTTTGGCCACTGATGC 777
2573 GAAGAGTCACTCACTGCAAAAAGCCCTGATGCTGGGAGGATTTGGGGGAGGAGAGAA 2632
778 GAAGAGTCACTCACTGCAAAAAGCCCTGATGCTGGGAGGATTTGGGGGAGGAGAGAA 837
2633 GGGGACGACAGAGATGATGGCTGGATGCATCATCTGACTCGATGACCTGAGTCTGG 2692
838 GGGGACGACAGAGATGATGGCTGGATGCATCATCTGACTCGATGACCTGAGTCTGA 897
2693 GTGACTCC 2701
898 GTGAACTCC 906

CC480290 907 bp DNA linear GSS 16-JUN-2003
CH240_307N7.T7 CHORI-240 Bos taurus genomic clone CH240_307N7,
genomic survey sequence.
CC480290
CC480290.1 GI:31758678
GSS.
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 907)
Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,
Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
Dalrymple, B.P. and Tellam, R.
Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other GSSs: CH240_307N7.FARBAC13p2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
Tel: 604-877-6065

Query Match		21.8%; Score 779.601; DB 1; Length 949;
Best Local Similarity		92.5%; Pred. No. 0.16;
Matches		882; Conservative 0; Mismatches 64; Indels 8; Gaps 6;
QY	1819	AGGTGAGGAGCAACAGTGTAGAACTGGACATGGAACACAGAGCTGGTTCCAAAGTAGGAAA 1878
DB	949	AGGTGAGGAGCAACAGTGTAGAACTGGACATGGAACACAGAGCTGGTTCCAAAGTAGGAAA 891
QY	1879	AGGATGATGTCAAGCTGTATATTTCTACCCGGCTTTGTTAACTTTCTATGAGAG-ACAT 1937
DB	890	AGGATGATGTCAAGCTGTATATTTCTACCCGGCTTTGTTAACTTTCTATGAGAGTACAT 831
QY	1938	CATGAGAAACGCTGGCTGGAGAGACACAGCTGGATCAAGATTCGCGGAGAAATAG 1997
DB	830	CATGAGAAATGCTGAGCTGGAGAGAAACAACTGGAATCAAGATTTGCCGGGAGAAATAT 771
QY	1998	CAATTAACCTCAGATATGAGATGATPACCACCTTTATGGCAGAAAGTGAAGAGGAACATAA 2057
DB	770	CAATTAACCTCAGATATGAGATGATPACCACCTTTATGGCAGAAAGTGAAGAGGAACTCGA 711
QY	2058	AAGCTCTTGTAGTGAAGTGAAGAGGAGAGTGAAGAGTGGCTTAAAGCTCAACATCA 2117
DB	710	AAGCTCTTGTAGTGAAGTGAAGAGGAGAGTGAAGAGTGGCTTAAAGCTCAACATCA 651
QY	2118	GAAACGAGATCATGGCATCTGGTCCCATCACTTTCATGGGAATAGATGGGGAACAGT 2177
DB	650	GAAATGAAGATCATGGCATTTAGTCCACACCTTCATGGGAATAGATGGGGAACAGT 591
QY	2178	GGAACAGTGTCACTTTATTTTGGGGGGCTCCAAATCACTGAGATGTGACTGC 2237
DB	590	GGAACAGTGTCACTTTATTTT--CTGGGCTCCAAATCACTGAGATGTGACTGC 533
QY	2238	AGCCATGAAATTAAGACACTTACTCTCTGGAGAGAAAGTTA--ACCACTTAGATAGC 2295
DB	532	AGCCATGAAATTAAGACGCTTACTCTCTGGAGAGAAAGTTATGACCAATCTAGATAGC 473
QY	2296	ATATTGAAAGCAGAGACATPACTCTGCCAACAAAGCCCATCTAGTCAAGGCTATGGTT 2355
DB	472	ATATTGAAAGCAGAGGACATTTCTGCCAACAAAGGTTCTGTAGTCAAGGCTATGGTT 413
QY	2356	TTTCCAGTGGTATGATGTGAGTGGAGTGGACTGTGAGAGAGCTGAGAGCTGAGCAAGA 2415
DB	412	TTTCCAGTGGTATGATGTGAGTGGAGTGGACTGTGAGAGAGCTGAGAGCTGAGCAAGA 353
QY	2416	ATTGATGCTTTTGAAGTGTGGTGTGAGAGAGCTCTTGAGAGTCCCTTGGAGTCAAGG 2475
DB	352	ATTGATGCTTTTGAAGTGTGGTGTGAGAGAGCTCTTGAGAGTCCCTTGGAGTCAAGG 293
QY	2476	AGATCCACAGTCCATCTGAAGAGATCAGCCCTGGGATTTCTTTGGAAGAAATGATG 2535
DB	292	AGATCCACAGTCCATCTGAAGAGATTAGCCCTTGGGATTTCTTTGGAAGAAATGATG 233
QY	2536	CTAAAGTGAATCTCAGTACTTTGGCCCTCTCATCAGAGAGCTGACTCACTCGGAAG 2595
DB	232	CTGAAGTGAATCTCAGTACTTTGGCCCTCTCATCAGAGAGCTGACTCACTCGGAAG 173
QY	2596	ACCTGATGCTGGAGGAGATTTGGGGCAGAGAGAGAGGGGACAGAGAGTGAATGG 2655
DB	172	ATTCTGATGCTGGAGGAGATTTGGGGCAGAGAGAGAGGGGACAGAGAGTGAATGG 113
QY	2656	CTGGATGGATCACTCACTGATGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2715
DB	112	CTGGATGGATCACTCACTGATGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 54
QY	2716	GACAGGAGGCTGTCTCTGGGCGATTTCAATGGGGTCAAAAGAGTTGGACAGA 2769
DB	53	AAGGGGAGGCTGTCTCTGGGCGATTTCAATGGGGT--GCAAGAGTGGACAGA 1
RESULT 26		
CC541222/c		
LOCUS		

DEFINITION

CH240_42205.T7 CHORI-240 Bos taurus genomic clone CH240_42205, genomic survey sequence.

ACCESSION

CC541222

VERSION

CC541222.1 GI:31859506

KEYWORDS

GSS.

SOURCE

Bos taurus (cow)

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 890)

AUTHORS

Holt,R., Scott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L., Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M., Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R., Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S., Schein,J., Marra,M., de Jong,P., Keele,J.W. and Kappes,S.M.

Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478 Unpublished (2003)

JOURNAL

Other_GSSs: CH240_42205.TARBAC13P2

COMMENT

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bccsc.ca

Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/ordering-information.html). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 422 row: 0 column: 5

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..890

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="bred: Hereford"

/db_xref="taxon:9913"

/clone="CH240_42205"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

Query Match

Best Local Similarity

21.8%; Score 779.1; DB 1; Length 890;

Matches 834; Conservative 0; Mismatches 54; Indels 5; Gaps 2;

QY

341

GCCTTGAGAACCCCATGAACAGTATGAAAGGCAAAATGATA--GGATCTGAAAGAGGA 398

DB

890

GCCTTGAGAACCCCATGAACAGTATGAAAGGCAAAATGATA--GGATCTGAAAGAGGA 831

QY

399

ACTCCAGGTCAGTAGTGTCCCATATGCTACTGGAGATCAGTGAGAAATCACTCCAG 458

DB

830

ATCCCGAGGTCAGTAGTGTCCCATATGCTACTGGAGATCAGTGAGAAATCACTCCAG 771

QY

459

AAAGATGAAGAGATGAGGCAAAAGCAAAAGATATCCAGCTGTGATCTGACTGTGA 518

DB

770

GAAGATGAAGGATGAGGCAAAAGCAAAAGATATCCAGCTGTGATCTGACTGTGA 711

QY

519

TATAAGCAAGTCCGATGTGTAAAGAGCAATATTCATAGGAACCTGGAATGTTCAGGTC 578

DB

710

TAGAAGGAAGTTCGAATGTGTAAAGAGCAAAATTCGTTAGGAACCTGGAATGTTCAGGCC 651

QY

579

CATGAATCAAGGCAAAATTCGAATGTGTAAAGAGATGGAAGAGTGAATCTCAACAT 638

Db 221 GAATAACAGGCAATTTGGCCCTGGAGTACGGAATGAAGCAGGCAAAAGGCTTAATAGAGT 162
 QY 1064 TTTCACAGAAATGCACTGGTCAATAGCAAAACACCTCTTCCAAACACACAGAGAAC 1123
 Db 161 TTTCACAGAGACACACCTGGTCAATAGCAAAACACCTCTTCCAAACACAGAGAAC 102
 QY 1124 TCTACATGGACATCACAGATGGTCAACACCGAAATCAGATTGATTATTTTGA 1183
 Db 101 TTTCACATGGACATCACAGATGGTCAACACCGAAATCAGATTGATTATTTTGA 42
 QY 1184 GCC-AAAGATGAGAGCTCTATACAGTCAGTCAGCAAAACAAG 1223
 Db 41 GCCGAAAGATGAGAGCTCTATACAGTCAGTCAGCAAAACAAG 1

RESULT 28
 CC511457/c
 LOCUS
 DEFINITION CH240_355A2.T7 CHORI-240 Bos taurus genomic clone CH240_355A2,
 Genomic survey sequence.

ACCESSION CC511457
 VERSION CC511457.1 GI:31829745

KEYWORDS GSS.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

1 (bases 1 to 945)
 Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B. P. and Tellam, R.

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 Unpublished (2003)

Other GSSs: CH240_355A2.TARBAC13P2
 Contact: Rob Holt

Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6

Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.

Plate: 355 row: A column: 2
 Seq primer: T7
 Class: BAC ends.

Location/Qualifiers
 1..945

/organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="Breed: Hereford"

/db_xref="taxon:9913"
 /clone="CH240_355A2"
 /sex="Male"

/cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull Li Domingo 9375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 21.8%; Score 778.1; DB 1; Length 945;
 Best Local Similarity 92.0%; Pred. No. 0.17;

Matches 873; Conservative 0; Mismatches 69; Indels 7; Gaps 5;
 QY 1815 ATCAGGTACGGAAGCAACAGTTAGAACTGGACATGGAACACAGACTGGTTCCAAGTAG 1874
 Db 945 ATCAGGTACGGAAGCAACAGTTAGAACTGGACATGGAACACAGACTGGTTCCAAGTAG 886
 QY 1875 GAAAGAGATATGTCAGGCTGTATTTGTCACCCCGCTTGTAACTTCTATGACAGAG- 1933
 Db 885 GAAAGAGATGTCAGGCTGTATTTGTCACCCCGCTTGTAACTTCTATGACAGAGT 826
 QY 1934 ACATCATGAGAAACGCTGGCTGGAAAGACCAAGCTGGAATCAAGATTGCGAGAGAA 1993
 Db 825 ACATCATGAGAAACGCTGGCTGGAAAGACCAAGCTGGAATCAAGATTGCGAGAGAA 766
 QY 1994 ATAGCAATTAACCTCAGATATGAGATGATACCAACCTTATGCGAGAAAGTG-AGAGGAAC 2053
 Db 765 ATATCATTAACCTCAGATATGAGATGATACCAACCTTATGCGAGAAAGTG-AGAGGAAC 707
 QY 2054 TAAAAAGCCTTTGATGAAAGTGAAAGAGGAGAGTGAATAAGTTGGCTTTAAAGCTCAACA 2113
 Db 706 TAAAAAGCCTTTGATGAAAGTGAAAGTGAGAGTGAATAAGTTGGCTTTAAAGCTCAACT 647
 QY 2114 TTCAGAAACGAGATCATGCGATCTGGTCCCATCACTTATGCGAAATAGATGGGAAA 2173
 Db 645 TTCAGAAATGAAGATCATGCGATCTGGTCCCATCACTTATGCGAAATAGATGGGAAA 587
 QY 2174 CAGTGGAAACAGTGTCAAGCTTTTATTTTGGGGGGCTCCAAAATCACTGCAAGTGGTGA 2233
 Db 586 CAATAGAAACAGTGTCAAGCTTTTATTTT-CTGGGCTCCAAAATCGCTACAGATGGTGA 529
 QY 2234 CTCGAGCCATGAAATTAAGACACCTTCTCTGGAAAGAAAGTTA--ACCAACTAGA 2291
 Db 528 CTCGAGCCATGAAATTAAGACACCTTCTCTGGAAAGAAAGTTAAGCAACACTAGA 469
 QY 2292 TAGCATATTGAAAGCAGAGACATTACCTTGGCAACAAAGCCCCATCTAGTCACAGGCTAT 2351
 Db 468 TAGCATATTCAAGACAGAGACATTACTTGGCCACCAAGTTTCTCTAGTCACAGGCTAT 409
 QY 2352 GCTTTTCCAGTGTCTATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2411
 Db 408 GCTTTTCTTGT 349
 QY 2412 AAGAATTCATGCTTTTGAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2471
 Db 348 ACGAATTCATGCTTTTGAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 289
 QY 2472 AAGGAGATCCAAACCAAGTCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAGGAAT 2531
 Db 288 AAGGAGATCCAAACCAAGTCCATTCTGAAGAGATCAGCCCTGGGATTTCTTTGGAGGAAT 229
 QY 2532 GATGCTTAAGCTGAACCTCCAGTACTTGGCCACTGATCAGAGAGCTGACTCACTGGA 2591
 Db 228 GATGCTTAAGCTGAACCTCCAGTACTTGGCCACTGATCAGAGAGCTGACTCACTGGA 169
 QY 2592 AAAGACCTGT 2651
 Db 168 AAAGACCTGT 110
 QY 2652 ATGCTGTGATGGATCACTGATCTGATGAGAGCTGATGAGTCTGGGTGAATCTCTGGAGTTGGT 2711
 Db 109 ATGCTGTGATGGATCACTGATCTGATGAGAGCTGATGAGTCTGGGTGAATCTCTGGAGTTGGT 50
 QY 2712 GATGACAGAGGAGGCTGTCTCTGGGGGATTCATGCGGTGTCACAAAGAGT 2760
 Db 49 GATGACAGAGGAGGCTGTCTCTGGGGTGTCTGATTCATGCGGTGTCACAAAGTGT 1

RESULT 29

CC495104

LOCUS

DEFINITION CH240_330L9.T7 CHORI-240 Bos taurus genomic clone CH240_330L9,

genomic survey sequence.

ACCESSION CC495104

CC495104 833 bp DNA linear GSS 17-JUN-2003

CH240_330L9.T7 CHORI-240 Bos taurus genomic clone CH240_330L9,

genomic survey sequence.

ACCESSION CC495104

VERSION	CC495104.1	GI:31808185
KEYWORDS	GSS.	
SOURCE	Bos taurus (cow)	
ORGANISM	Bos taurus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.	
AUTHORS	1 (bases 1 to 833) Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L., Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M., Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S., Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W., Dalrymple, B. P. and Tellam, R.	
TITLE	Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398	
JOURNAL	Unpublished (2003)	
COMMENT	Other GSSs: CH240 330L9.TARBAC13P2 Contact: Rob Holt Sequencing The British Columbia Cancer Agency Genome Science Centre 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6 Tel: 604-877-6085 Fax: 604-877-6276 Email: rholt@bcgsc.ca Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering/information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada. Plate: 330 Row: L Column: 9 Seq primer: T7 Class: BAC ends.	
FEATURES	Location/Qualifiers	
source	1..833	
	/organism="Bos taurus"	
	/mol_type="genomic DNA"	
	/strain="bred: Hereford"	
	/db_xref="taxon:9913"	
	/clone="CH240_330L9"	
	/sex="Male"	
	/cell_type="Blood"	
	/clone_lib="CHORI-240"	
	/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull LI Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"	
Query Match	21.7%; Score 775.4; DB 1; Length 833;	
Best Local Similarity	95.7%; Pred. No. 0.19;	
Matches 797; Conservative	0; Mismatches 36; Indels 0; Gaps 0;	
QY	352	CCCATGACAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGACTCCCGAGTCA 411
DB	1	CCCATGAATAGTATGAAAGGCAAAATGATAGGATCTGAAAGAGGACTCCCGAGTCC 60
QY	412	GTAGGTGCCATATGCTACTGCGATCAGTGGAGAAATACTCCAGAAAGAAATGAAG 471
DB	61	GTAGGGGCCATAGCTACTGCGATCAGTGGAGAAATACTCCAGAAAGAAATGAAGGG 120
QY	472	ATGGAGCCCAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGTGGATATAGCAAGT 531
DB	121	ATGGAGCCCAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGTGGATATAGCAAGT 180
QY	532	CGATCTGTAAAGACATATTGCGATAGCAACTGGAATGTGAGTCCATGAATCAAGGC 591
DB	181	CGATCTGTAAAGACATATTGCGATAGCAACTGGAATGTGAGTCCATGAATCAAGGC 240
QY	592	AAATTGGAAGTGGTCAAAACAGAGATGCGAAGAGTCAATGTCAACATCTTAGGAATCAG 651
DB	241	AAATTGGAAGTGGTCAAAACAGAGATGCGAAGAGTCAATGTCAACATCTTAGGAATCAG 300
QY	652	GAACTAAATGCACTGGATGGTGAATTTAACTCAGATGACCATATATATCTACTACTGC 711
DB	301	AAACTGAAATGCACTGGATGGTGAATTTAACTCAGATGACCATATATCTACTACTGT 360
QY	712	GGCCAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCCGAAT 771
DB	361	GGACAGGAATCCCTCAGAGAAATGGAGTAGCCATCATGGTCAACAAAAGAGTCCGAAT 420
QY	772	GCAGTACTTGGATGAGTCTCAAAAACACAGAAATGATCTCTGTTGTTTCCAAAGGCAA 831
DB	421	GCAGTACTTGGATGAGTCTCAAAAACACAGAAATGATCTCTGTTGTTTCCAAAGGCAA 480
QY	832	CAATTCAATATCACAGTATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAGAGCTGAA 891
DB	481	CAATTCAATATCACAGTATCCAAAGTCTATGCCCCCAACAGTAATGCTGAAGAGCTGAA 540
QY	892	GTGAAACGGTCTCTATGAGACCTACAAGACCTTTTAGAACTAACACCCCAAAAAGATGTC 951
DB	541	GTGAAACGGTCTCTATGAGACCTACAAGACCTTTTAGAACTAACACCCCAAAAAGATGTC 600
QY	952	CTTCTCATTTATAGGGGACTGGAAATGCAAAAGTAGGAAGCAAAAGAACACCTCGAGTAACA 1011
DB	601	CTTCTCATTTATAGGGGACTGGAAATGCAAAAGTAGGAAGCAAAAGAACACCTCGAGTAACA 660
QY	1012	GGCAAAATTTGGCTTTGGAATACGGAATGAAGCAGGCAAAAGACTAATAGAGTTTGGCAA 1071
DB	661	GGCAAAATTTGGCTTTGGAATACGGAATGAAGCAGGCAAAAGACTAATAGAGTTTGGCAA 720
QY	1072	GAAATGCACTGGTCTATGCAACACCCCTCTTCCACACACACAGAGAGACTCTACACA 1131
DB	721	GAAATGCACTGGTCTATGCAACACCCCTCTTCCACACACACAGAGAGACTCTACACA 780
QY	1132	TGGACATCACCAGATGGTCAACACCCGAAATCAGATTGATTATATCTTTTCAG 1184
DB	781	TGGACATCACCAGATGGTCAACACCCGAAATCAGATTGATTATATCTTTTCAG 833
RESULT 30	CC539858 887 bp DNA linear GSS 17-JUN-2003	
LOCUS	CH240_420A14.T7 CHORI-240 Bos taurus genomic clone CH240_420A14,	
DEFINITION	genomic survey sequence.	
ACCESSION	CC539858	
VERSION	CC539858.1	
KEYWORDS	GSS.	
SOURCE	Bos taurus (cow)	
ORGANISM	Bos taurus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.	
REFERENCE	1 (bases 1 to 887)	
AUTHORS	Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L., Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M., Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S., Schein, J., Marra, M., de Jong, P., Keels, J. W. and Kappes, S. M.	
TITLE	Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478	
JOURNAL	Unpublished (2003)	
COMMENT	Other GSSs: CH240_420A14.TARBAC13P2 Contact: Rob Holt Sequencing The British Columbia Cancer Agency Genome Science Centre 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6 Tel: 604-877-6085 Fax: 604-877-6276 Email: rholt@bcgsc.ca Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering/information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the	


```

Db      836  CCGTGGACTGGAGGAAACACACAGCTGGAAATCAGATTCGCAAGGAAAT--CAATCACT 779
QY      2007  CAGATATGAGATGATACACCCCTTATGGCAGAAAGTGAAGAGGAATCAAAAACCTCTT 2066
Db      778  CAGATATGAGATGATACACCCCTTATGGCAGAAAGTGAAGAGGAATCAAAAACCTCTT 719
QY      2067  GATGAAGGTGAAGAGGAGAGTGAAGAGTGGCTTAAAGCTCAACATTCAGAAAGCGAA 2126
Db      718  GATGAAGTGAAGAGTGGAGAGTGAAGAGTGGCTTAAAGCTCAACATTCAGAAAGCGAA 659
QY      2127  GATCATGGCATCTGTGCCCATCACTCATGGGAATAGATGGGAAACAGTGGAAACAGT 2186
Db      658  GATCATGGCATCTGTGCCCATCACTCATGGGAATAGATGGGAAATAGTGGAAACAT 599
QY      2187  GTCAGACTTTATTTTGGGGGGCTCCAAATCACTGAGATGCTGAGCTGAGCCATGAA 2246
Db      598  GTCAGACTTTATTTT--CTGGGCTCCAAATCACTGAGATGCTGAGCTGAGCCATGAA 541
QY      2247  ATTAAGACACTTACTCTCTTGGAAAGAAAGTTA--ACCAACCTAGATAGCATATGAAA 2304
Db      540  ATTAAGACAGCTTACTCTCTTGGAAAGAAAGTTATGACCAACCTAGGTAGCATATGAAA 481
QY      2305  AGCAGAGCATTAACCTTGTCCACAAAGCCCATCTAGTCAAGGCTATGTTTCCAGTG 2364
Db      480  AGCAGAGCATTAACCTTGTCCACAAAGAGTCTGCTAGTCAAGGCTATGTTTCCAGTG 421
QY      2365  GTCATGATGAGATGAGAGTTGAGCTGTGAAGAAAGCTGAGCACTGAAGAAATGATGCT 2424
Db      420  GTCATGATGAGATGAGAGTTGAGCTGTGAAGAAAGCTGAGCACTGAAGAAATGATGCT 361
QY      2425  TTTGAAGCTGCTGTTGGAGAGAGCTCTTGAAGAGTCCCTTGAAGTGGCTGAGCACTGAAG 2484
Db      360  TTTGAAGCTGCTGTTGGAGAGAGCTCTTGAAGAGTCCCTTGAAGTGGCTGAGCACTGAAG 301
QY      2485  CAGTCCATCTCAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATCTAAAGCTG 2544
Db      300  CAGTCCATCTCAAGAGATCAGCCCTGGGATTTCTTTGGAAGGAATGATCTAAAGCTG 241
QY      2545  NAACTCCAGTACTTTGGCCACTGATCAGAGAGCTGACTCAGTGGAAAGACCTCATG 2604
Db      240  NAACTCCAGTACTTTGGCCACTGATCAGAGAGCTGACTCAGTGGAAAGAGTCTCATG 181
QY      2605  CTGGAGGAGTGGGGGAGGAGAGAGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2664
Db      180  CTGGAGGAGTGGGGGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 121
QY      2665  ATCACTGACTGAGAGCTGAGTCTGGTGAACCTCTGGAGTTGCTGATGGACAGGAG 2724
Db      120  ATCACTGACTGAGAGCTGAGTCTGGTGAACCTCTGGAGTTGCTGATGGAGGAGGAG 61
QY      2725  GCCTGCTCTGCGCGATTTCATGGGCTCAAGAGAGTTGGACACGACTGAGCAACTGAACT 2784
Db      60  GCCTGCTGCTGCGATTTCATGGGCTCAAGAGAGTTGGACACGACTGAGCAACTGATCT 1

RESULT 32
CC508740
LOCUS
DEFINITION
CH240_351A11.TARBAC13P2 CHORI-240 Bos taurus genomic clone
CH240_351A11, genomic survey sequence.
ACCESSION
CC508740
VERSION
1
KEYWORDS
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE
1 (bases 1 to 827)
Holt,R., Stott,J., Yang,G., Barber,S., Smallos,D., Prabhu,A.-L.,
Tsal,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,

```

Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
Dallrymple, B. P., and Tellam, R.
Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
Unpublished (2003)
Other GSSs: CH240_351A11.T7
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 351 row: A column: 11
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1..827
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_351A11"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match 21.7%; Score 174.2; DB 1; Length 827;
Best Local Similarity 96.0%; Pred. No. 0.2;
Matches 794; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 438 TCAGTGGAGAAATAACTCCAGAAAGATGAAGAGATGGAGCCAAAGCAAAAGAAATACCC 497
Db 1 TCAGTGGAGAAATAACTCCAGAAAGATCTGTTGGATGGAGCCAAAGCAAAAGCAATACCC 60
QY 498 AGCTGTGATGTGCTGTGATATAGCAAGTCCGATGCTGTAAGAGCAATATTCAT 557
Db 61 AGCTGTGATGTGCTGTGATAGAGCAAGTCCGATGCTGTAAGAGCAATATTCAT 120
QY 558 AGGAACCTGGAATGTCCAGTCCATGAATCAAGGCAAAATTTGGAAGTGTCAAAAGAGAT 617
Db 121 AGGAACCTGGAATGTCCAGTCCATGAATCAAGGCAAAATTTGGAAGTGTCAAAAGAGAT 180
QY 618 GGCAGAGTGAATGTCAACATTTAGGAATCAGGAACTAAATGGACTGGATGGGTGA 677
Db 181 GGCAGAGTGAATGTGCAATTTAGGAATCAGGAACTAAATGGACTGGATGGGTGA 240
QY 678 ATTTAACTCAGATGACCAATATATCTACTCTGCGGAGCAAAATCCCTCAGAGAAATGG 737
Db 241 ATTTAACTCAGATGACCAATATATCTACTCTGCGGAGCAAAATCCCTCAGAGAAATGG 300
QY 738 AGTAGCCATGTGTCACAAAGAGTCCGAAATGCACTACTGATGCTGAGTCTCAAAAA 797
Db 301 AGTAGCCATGTGTCACAAAGAGTCCGAAATGCACTACTGATGCTGAGTCTCAAAAA 360
QY 798 CGACAGATGATCTCTGTTGTTTCCAGGCAAAACCAATTCATATCACAGTAATCCAAAT 857
Db 361 TGACAGATGATCTCTGTTGTTTCCAGGCAAAACCAATTCATATCACAGTAATCCAAAT 420
QY 858 CTATGCCCCCAACAGTAAATGCTGAAGAGTGAAGTGAACCGTCTTATGAAGACCTACA 917
Db 421 CTATGCCCCCAACAGTAAATGCTGAAGAGTGAAGTGAACCGTCTTATGAAGACCTACA 480

QY 918 AGACCTTTTAGAAGTAAACCCCAAAAAGATGCTCTTCTCATTTATAGGGAGTGAATGC 977
 Db |||||
 QY 481 AGACCTTTTAGAAGTAAACCCCAAAAAGATGCTCTTCTCATTTATAGGGAGTGAATGC 540
 Db |||||
 QY 978 AAAAGTAGGAAGCAAGAAACACCTGGAGTAACAGGCAAAATTTGGCCCTTGGAAATACGGAA 1037
 Db |||||
 QY 541 AAAAGTAGGAAGTCAAGAAACACCTGGAGTAACAGGCAAAATTTGGCCCTTGGAAATACGGAA 600
 Db |||||
 QY 1038 TGAAGCAGGGCAAGAGCTAATAGATTTTGGCAAGAAATGCACTGGTTCATAGCAACAC 1097
 Db |||||
 QY 601 TGAAGCAGGGCAAGAGCTAATAGATTTTGGCAAGAAATGCACTGGTTCATAGCAACAC 660
 Db |||||
 QY 1098 CCTCTTCCAAACACACAGAGAGACTCTACACATGGACATCACACAGATGCTCAACACCG 1157
 Db |||||
 QY 661 CCTCTTCCAAACACATAGAGAGACTCTACACATGGACATCACACAGATGCTCAACACCA 720
 Db |||||
 QY 1158 AAATCAGATGATTAATATCTTTTGCAGCAAGATGGAGAGCTCTATACAGTCAACAAA 1217
 Db |||||
 QY 721 AAATCAGAGCTGATTAATATCTTTTGCAGCAAGATGGAGAGCTCTATACAGTCAACAAA 780
 Db |||||
 QY 1218 AACAGACACAGAGCTTACTGTGGCTCAGATCATCACTCTCTTTATG 1264
 Db |||||
 QY 781 AACAGACACAGAGCTGACTGTGGCTCAGATCATCACTCTCTTTATG 827
 Db |||||

RESULT 33

CC512039
 LOCUS CH240_356E9.T7 CHORI-240 Bos taurus genomic clone CH240_356E9,
 genomic survey sequence.

ACCESSION CC512039

VERSION CC512039.1 GI:31830327

KEYWORDS GSS.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

REFERENCE

AUTHORS Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
 Tsai,M., Cloutier,A., Lee,D., Girm,N., Olson,T., Mayo,M.,
 Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
 Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
 Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
 Dalrymple,B.P. and Tellam,R.

TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

JOURNAL Unpublished (2003)

COMMENT Other GSSs: CH240_356E9.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bccsc.ca

Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.choi.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/ordering/information.htm). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 356 row: E column: 9

Seq primer: T7

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..898

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clone="CH240_356E9"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/notes="Vector: pTARBAC1.3; Site_1: Mbol; Site_2: Mbol;

Hereford bull Li Domino 99375; CHORI-240 Bovine BAC

Library (Male) produced by Pieter de Jong"

Query Match 21.7%; Score 774.2; DB 1; Length 898;
 Best Local Similarity 94.0%; Pred. No. 0.18;
 Matches 846; Conservative 0; Mismatches 48; Indels 6; Gaps 4;

QY 1796 TGACTCTTGAATAATTTGTATGCAAGTCAAGGCTGTATATGTCCACCGGCTTG 1915
 Db 1 TGCTCTTGAATAATTTGTATGCAAGTCAAGGCTGTATATGTCCACCGGCTTG 1855
 QY 1856 ACAGACTGGTTCACAGTAGGAAAGGAGTATGTCAAGGCTGTATATGTCCACCGGCTTG 1915
 Db 61 ACAGACTGGTTCACAGTAGGAAAGGAGTATGTCAAGGCTGTATATGTCCACCGGCTTG 120
 QY 1916 TTTAACTTCTATGCGAGAG-ACATCATAGAAACGCTGGCTGGAAGCAAGCAAGCTGGA 1974
 Db 121 TTTAACTTCTATGCGAGTACATCATGACACACGCTGGACTGGAAGCAAGCAAGCTGGA 180
 QY 1975 ATCAAGATTGCCGGAGAAATAGCAATACCTCAGATATGCAAGTATGACACACCTTATG 2034
 Db 181 ATCAAGATTGCCAGAGAAATATCAATACCTCAGATATGCAAGTATGACACACCTTATG 240
 QY 2035 GCAGAAAGTGAAGAGGAACTAAAAAGCCTCTTGTATGAAAGGTGAAAGAGAGAGTGAAGAA 2094
 Db 241 GCAGAAAGTGAAGAGGAACTAAAAAGCCTCTTGTATGAAAGTGAAGAGTGAAGAGTGAAGAA 300
 QY 2095 GTTGCTTTAAAGCTCAACATTCAGAAACAGAGATCAAGGATCTGGTCCCATCACTTCA 2154
 Db 301 ATGGCTTTAAAGCTCAACATTCAGAAACAGAGATCAAGGATCTGGTCCCATCACTTCA 360
 QY 2155 TGGGAAATAGATGGGAAACAGTGTGAAACAGTGTGAGACTTTATTTTGGGGGCTCCA 2214
 Db 361 TGGGAAATAGATGGGAAACAGTGTGAAACAGTGTGAGACTTTATTTT-CTGGGCTCCA 418
 QY 2215 AAATCACTGCGAGATGGTGAAGTGTGAGCCATGAAATTAAGACACTTACTCTTGGAGAA 2274
 Db 419 AAATCACTGCGAGATGGTGAAGTGTGAGCCATGAAATTAAGACACTTACTCTTGGAGAA 478
 QY 2275 AAGT--TAACCAACCTAGATAGCATATTGAAAGCAGAGACATTACCTTGGCCCAAGC 2332
 Db 479 AAGTTATTAACCAACCTAGATAGCATATTGAAAGCAGAGACATTACTTTGCCAACAAGG 538
 QY 2333 CCATCTAGTCAAGGCTATGTTTTTCCAGTGGTCAATGTATGAGTGTGAGAGTTGGAGCTG 2392
 Db 539 TTCTCTAGTCAAGGCTATGTTTTTCCAGTGGTCAATGTATGAGTGTGAGAGTTGGAGCTG 598
 QY 2393 TCAAGAAAGCTGAGCAGCTGAAGAAATGATGCTTTTGAAGTGTGAGTGTGAGAGAGCTCT 2452
 Db 599 TGAAGAGGCTGAGCAGCCGAAAGAAATGATGCTTTTGAAGTGTGAGTGTGAGAGAGCTCT 658
 QY 2453 TGAGAGTCCCTTGGAGTGCAGAGAGATCCCAACCAAGTCCATTCTGAAAGAGATCAGCCCTG 2512
 Db 659 TGAGAGTCCCTTGGAGTGCAGAGAGATCCCAACCAAGTCCATTCTGAAAGAGATCAGCCCTG 718
 QY 2513 GGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAGTCCAGTACTTTGGCCACCTGATCA 2572
 Db 719 GGATTTCTTTGGAAGGAATGATGCTAAAGCTGAAGTCCAGTACTTTGGCCACCTGATGC 778
 QY 2573 GAAGAGTGTGACTCACTCTGAAAGAGACCTGATGCTGGGAGGAGATTGGGGGAGGA-GGAGA 2631
 Db 779 GAAGAGTGTGACTCACTCTGAAAGAGACCTGATGCTGGGAGGAGATTGGGGGAGGAGGA 838
 QY 2632 AGGGGACAGCAGAGATGAGATGGCTGGATGGCATCACTGATCGATGAGAGTGGAGTCTG 2691
 Db 839 AGGGGACAGCAGAGATGAGATGGCTGGATGGCATCACTGATCGATGAGAGTGGAGTCTG 898

Db	181	ATCAAGATTGCCAGGAGAAATATCAATAACCTCAGATATGTCAGATGACACACCCCTTATG	240						
QY	2035	GCAGAAAGTGAAGAGGAACCTTAAAGAGCTCTTTCATGAAGGTGAAGAGGAGAGTGA	2094						
Db	241	GCAGAAAGTGAAGAGGAACCTTAAAGAGCTCTTTCATGAAGGTGAAGAGGAGAGTGA	300						
QY	2095	GTTGGCTTAAAGCTCAACATTCAGAAAACGAGAGATCATGGCATCTGGTCCCATCTCA	2154						
Db	301	GTTGGCTTAAAGCTCAACATTCAGAAAACGAGAGATCATGGCATCTGGTCCCATCTCA	360						
QY	2155	TGGGAATATAGATGGGAAACAGTGGAAACAGTGTCTAGACTTTATTTTGGGGGCTCCA	2214						
Db	361	TGGGAATATAGATGGGAAACAGTGGAAACAGTGTCTAGACTTTATTTTCTGGGGCTCCA	418						
QY	2215	AAATCACTGCAGATGGTGCAGCCATGAAATTTAAAGACACTTACTCTCTTGGAGAA	2274						
Db	419	AAATCACTGCAGATGGTGCAGCCATGAAATTTAAAGACCGCTTACTCTCTTGGAGAA	478						
QY	2275	AAGTTA--ACCAACTAGATAGCATTTTAAAGACAGACATTAACCTTGGCCAAACAGC	2332						
Db	479	AAGTTATGACCAACCTAGATAGCATTTTAAAGACAGACATTAACCTTGGCCAAACAGG	538						
QY	2333	CCCATCTAGTCAAGGCTATGTTTTTCCAGTGGTCATGTATGGATGTGAGAGTTGGACTG	2392						
Db	539	TTTCGCTAGTCAAGGCTATGTTTTTCCAGTGGTCATGTATGGATGTGAGAGTTGGACTG	598						
QY	2393	TGAAGAAAGCTGAGCACTGAAGAATGTCTTTTGAAGTGTGGTGGAGAGACTCT	2452						
Db	599	TAAAGAGGCTGAGCGCGGAAGATTTGATGCTTTTGAACCTGTGGTGTGGAGAGAGTCT	658						
QY	2453	TCAGAGTCCCTTGGACTGCAAGGAGATCCAACAGTCCATCTTGAAGGAGATCAGCCCTG	2512						
Db	659	TGAGAGTCCCTTGGACTGCAAGGAGATCCAACAGTCCATCTTGAAGGAGATCAGCCCTG	718						
QY	2513	GGATTCTTTTGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTGATCA	2572						
Db	719	GGATTCTTTTGAAGGAATGATGCTAAAGCTGAAACTCCAGTACTTTGGCCACCTCATGT	778						
QY	2573	GGAAGAGTCACTTCTGAAAAGACCCGTGCTGGAGGAGATTTGGGGGAGAGAGAGAA	2632						
Db	779	GGAAGAGTTGAATTCATTTGAAAAGACTCTGATGCTGGGAGGATTTGGGGGAGAGAGAA	838						
QY	2633	GGGGACGACAGAGGATGAGATGGCTGGATGGCATCACTGA	2672						
Db	839	GGGGACGACAGAGGATGAGATGGCTGGATGGCATCACTGA	878						
RESULT 35	CC536979/c	915 bp	DNA	linear	GSS 17-JUN-2003				
LOCUS	CH240_416G6.T7	CHORI-240	Bos taurus	genomic clone	CH240_416G6,				
DEFINITION	genomic survey sequence.								
ACCESSION	CC536979								
VERSION	CC536979.1	GI:31855267							
KEYWORDS	GSS								
SOURCE	Bos taurus (cow)								
ORGANISM	Bos taurus								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.								
AUTHORS	1. (bases 1 to 915) Holt,R., Scott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L., Tsai,M., Cloutier,A., Lee,D., Girm,N., Olson,T., Mayo,M., Chiu,R., Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R., Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S., Schein,J., Marra,M., de Jong,P., Keelle,J.W. and Kappes,S.M.								
TITLE	Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478								
JOURNAL	Unpublished (2003)								
COMMENT	Other_GSSs: CH240_416G6.TARBAC13P2 Contact: Rob Holt Sequencing The British Columbia Cancer Agency Genome Science Centre 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6								

RESULT 34	CC581072	881 bp	DNA	linear	GSS 18-JUN-2003				
LOCUS	CH240_376A7.T7	CHORI-240	Bos taurus	genomic clone	CH240_376A7,				
DEFINITION	genomic survey sequence.								
ACCESSION	CC581072								
VERSION	CC581072.1	GI:31928363							
KEYWORDS	GSS								
SOURCE	Bos taurus (cow)								
ORGANISM	Bos taurus								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.								
AUTHORS	1. (bases 1 to 881) Holt,R., Scott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L., Tsai,M., Cloutier,A., Lee,D., Girm,N., Olson,T., Mayo,M., Chiu,R., Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R., Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S., Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W., Dalrymple,B.P. and Tellam,R.								
TITLE	Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398								
JOURNAL	Unpublished (2003)								
COMMENT	Other_GSSs: CH240_376A7.TARBAC13P2 Contact: Rob Holt Sequencing The British Columbia Cancer Agency Genome Science Centre 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6 Tel: 604-877-6085 Fax: 604-877-6276 Email: rholt@bcgsc.ca Clones are derived from the bovine BAC library CHORI-240 (http://www.chori.org/bacpac/bovine240.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering/information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada. Plate: 376 row: A column: 7 Seq primer: T7 Class: BAC ends.								
FEATURES	Location/Qualifiers								
source	1..881								
	/organism="Bos taurus"								
	/mol_type="genomic DNA"								
	/strain="breed: Hereford"								
	/db_xref="taxon:9913"								
	/clone="CH240_376A7"								
	/sex="Male"								
	/cell_type="Blood"								
	/clone_lib="CHORI-240"								
	/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"								
Query Match	21.7%; Score 773.7; DB 1; Length 881;								
Best Local Similarity	94.5%; Pred. No. 0.19;								
Matches 832; Conservative	0; Mismatches 43; Indels 5; Gaps 3;								
QY	1796	TGACTCTGAAAATTTGTCAGGTCAGGAGCAACAGTAGAAGTGGACATGGAAACA	1855						
Db	1	TGCCTCTTGAGAAATTTGTCAGGTCAGGAGCAACAGTAGAAGTGGACATGGAAACA	60						
QY	1856	ACAGACTGGTTCAGAGTAAAGAGTATGTCAAGGCTGTATTTCTACCGGCTTG	1915						
Db	61	ACAGACTGGTTCAGAGTAAAGAGTATGTCAAGGCTGTATTTCTACCGCTGTTTA	120						
QY	1916	TTTAACTTCTATGACAG-ACATCATGAGAAACGCTGGCTGGAGAGCAACAGCTGGA	1974						
Db	121	TTTAACTTCTATGACAGATCATCATGAGAAACCGCTGACTTGGAGAGAACACAGCTGGA	180						
QY	1975	ATCAAGATTGCCGGAGAAATAGCAATAACCTCAGATATGAGATGATACACCCCTTATG	2034						

Tel: 604-877-6085
Fax: 604-877-6276

Email: rholt@bcgs.ca

Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orderinginformation.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.

Plate: 416 row: G column: 6

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

FEATURES

source

1..915

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clones="CH240_416G6"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull Li Domino 99375; CHORI-240 Bovine BAC

library (Male) produced by Pieter de Jong"

Query Match

Best Local Similarity 93.8%; Score 771.3; DB 1; Length 915;

Matches 834; Conservative 0; Mismatches 52; Indels 3; Gaps 3;

QY 712 GGGCAGGATCCCTCAGAGAAATGGAGTAGCCATCATGTGTCAAC-AAAAGAGTCCGAAA 770
DB 913 GCGAGGGAATCCCTTAGAAGAAATGGAGTAGCCATCATGTGTCAACAAAAGAGTCCGAAA 854
QY 771 TGCAGTACTTGGATCGAGTCTCAAAAACGACAGATGATCTCTGTTTGTTCACAGGCA 830
DB 853 TGCAGTACTTGGATCGAGTCTCAAAAACGACAGATGATCTCTGTTTGTTCACAGGCA 794
QY 831 ACCATTCAATATCAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAGAAGCTGA 890
DB 793 ACCATTCAATATCAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAGAAGCTGA 734
QY 891 AGTTGAAGGTCCTATCAGAGCTTACAGACCTTTTAGACTTACACCCCAAAAAGATGT 950
DB 733 AGTTGAAGGTCCTATCAGAGCTTACAGACCTTTTAGACTTACACCCCAAAAAGATGT 674
QY 951 CCTTCTCATTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAACACCTGGAGTAAC 1010
DB 673 CCTTCTCATTATAGGGGACTGGAATGCAAAAGTAGGAAGCAAAACACCTGGAGTAAC 614
QY 1011 AGCCAAATTTGGCTTGGATAGGATAGGATGAGGCGCAAAAGCTAATAGTGTTCGCA 1070
DB 613 AGCCAAATTTGGCTTGGATAGGATAGGATGAGGCGCAAAAGCTAATAGTGTTCGCA 554
QY 1071 AGAAATGCACTGGTTCATAGCAAAACCCCTTCTTCCCAACCAAGAGAGACTCTACAC 1130
DB 553 AGAAGAGCACTGGTTCATAGCAAAACCCCTTCTTCCCAACCAAGAGAGACTCTACAC 494
QY 1131 ATGACATCACCAGATGGTCAACCCGAAATCAGATTGATTATTTCTTTGAGGCAAG 1190
DB 493 ATGACATCACCAGATGGTCAACCCGAAATCAGATTGATTATTTCTTTGAGGCAAG 434
QY 1191 ATGAGAGGCTCTATCAGTCAGCAAAAACAGACCCAGAGGCTTACTGTGGCTCAGATCA 1250
DB 433 ATGAGAGGCTCTATCAGTCAGTCGCAAAAACAGACCCAGAGGCTTACTGTGGCTCAGATCA 374
QY 1251 TGAATCTCTTATTTGCAAAATTCAGATTAAATTTGAAGAAGTAGGGAAGAA-CCACTAGAT 1309
DB 373 TGAGTCTCTTATTTG-CAAAATCAGACTTAAATTTGAAGAAGTAGGGAAGAAACCCCTGGAC 315
QY 1310 CACTCAGGTAAGACCTAAATCCATCCCTTATCATTTATATACAGTGGAGTGAAGATAGAT 1369

DB 314 CATTGGGTATGACCTTAATCAATCCCTTATGATTATACAGTGGAGTGACAAATAGAT 255
QY 1370 TTAAGGCCCTAGATCTGATAGACAGAGTAGTACCTAATGAATGAGTGGACAGAGTTTCATGACA 1429
DB 254 TTAAGGGACTAGATCTGATAGATAGAGTGCCTGATGAATGAGTGGACAGAGTTTCATGACA 195
QY 1430 TTTACAGAGACAGAGTGGATCGAGACCATCCCATGGAAGAAATGCAAAAAGCAAAAT 1489
DB 194 TTTACAGAGACAGAGTGGATCGAGACCATCCCATGGAAGAAATGCAAAAAGCAAAAT 135
QY 1490 GGCTGTCTGGGAGGCTTTACAAATAGCTGTGAAAAGAGAGAAAGTGAAGCAAGGAA 1549
DB 134 GGCTGTCTGGGAGGCTTTACAAATAGCTGTGAAAAGAGAGAGAGGAAAGCAAGGAG 75
QY 1550 AAAAGGAAGATAAAGCATCTGAATGACAGAGTTCCAAAGAACTTCCAG 1598
DB 74 AAAAGGAAGATAAAGCATCTGAATGACAGAGTTCCAAAGAACTTCCAG 26

RESULT 36

CC505149/c

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CC505149 846 bp DNA linear GSS 17-JUN-2003
CH240_345P10.77 CHORI-240 Bos taurus genomic clone CH240_345P10,
genomic survey sequence.

CC505149

CC505149.1 GI:31823442

GSS.

Bos taurus (cow)

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

1 (bases 1 to 846)

Holt, R., Stott, J., Yang, G., Barber, S., Smal, D., Prabhu, A., L.,

Tsai, M., Cloutier, A., Lee, D., Gilm, N., Olson, T., Mayo, M.,

Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,

Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,

Schein, J., Marra, M., de Jong, P., McWilliam, S., Barrie, W.,

Dalrymple, B.P., and Tellam, R.

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

Unpublished (2003)

Other_GSS: CH240_345P10.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6

Tel: 604-877-6085

Fax: 604-877-6276

Email: rholt@bcgs.ca

Clones are derived from the bovine BAC library CHORI-240

(http://www.chori.org/bacpac/bovine240.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/orderinginformation.htm). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 345 row: P column: 10

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..846

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="taxon:9913"

/clones="CH240_345P10"

/sex="Male"

/cell_type="Blood"

/clone_lib="CHORI-240"

/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

Hereford bull Li Domino 99375; CHORI-240 Bovine BAC

FEATURES

source

library (Male) produced by Pieter de Jonga"

Query Match	21.6%	Score	770.7	DB	1	Length	846
Best Local Similarity	95.8%	Pred. No.	0.2				
Matches	801	Conservative	0	Mismatches	34	Indels	1
Gaps	1						
355	ATGAACAGTATG-AAAAGGCAAAATGATAGGATACTGAAGAGGAACTCCCAAGTCAAGT	413					
846	ATGAACAGTATGAAAAGGCAAAATGATAGGATCTGAAGAGAACTCCCAAGTCAAGT	787					
414	AGGTGCCCATATGCTATCTGGAGATCAGTGGAGAAATAACTCCAGAAAGATGAAGAGAT	473					
786	AGGTGCCCAATGTGCTACTGGAGTCAAGTGGAGAAATAACTCCAGAAAGATGAAGGAT	727					
474	GGAGCCAAAGCAAAAGATACCCAGCTGTGCATGTGACTGGTGATATAAGCAAGTCCG	533					
726	GGAGCCAAAGCAAAACAATACCCAGCTGTGCATGTGACTGGTGATAGAAGCAAGTCTG	667					
534	ATGCTGTAAGAGACAATATTGCAATAGGAACCTGGAAATGTCAGGTCCATGAATCAAGGCA	593					
666	ATGCTGTAAGAGACAGTATTGCAATAGGAACCTGGAAATGTCAGGTCCATGAGTCAAGGCA	607					
594	ATTGGAAAGTGTCAACCAAGAGATGCCAGAGTGAATGTCACAACTCTAGCAATCAGCGA	653					
606	ATTGGAAAGTGTCAACCAAGAGTGAATGCGAAGTGAATGTCACAACTCTAGCAATCAGTGA	547					
854	ACTAAATGGAAGTGGAAATGGGTGCAATTTAACTCAGATGACCAATATATCTACTACTCGG	713					
546	ACTCAATGGACTGGAAATGGGTGCAATTTAACTCAGATGACCAATATATCTACTACTCGG	487					
714	GCAGAAATCCCTCAGAAAGAAATGGAGTAGCCATCATGTCACAAAGAGTCCGAATGTC	773					
486	GCAGAAATCCCTCAGAAAGAAATGGAGTAGCCATCATGTCACAAAGAGTCCGAATGTC	427					
774	AGTACTTGGATGTCAGTCTCAAAACACAGCAATGATCTCTGTCTGTTTCCAAAGGCAAAAC	833					
426	AGTACTTGGATGTCAAATCTCAAAATGACAGATGATCTCTTTTCATTTTCCAAAGCAAAAC	367					
834	ATTCAATATCACAGTAATCAAGTCTATGCCCAACACAGTAATGCTGAAGAGTCTGAAGT	893					
366	ATTCAATATCACAGTAATCAAGTCTATGCCCAACACAGTAATGCTGAAGAGTCTGAAGT	307					
894	TGACGGTCTTATCAAGACTTCAAGACCTTTTAGAACTTAACACCAAAAAAGATGTCTCT	953					
306	TGACGGTCTTATGAAGACCTTCAAGACCTTTTAGAACTTAACACCAAAAAAGATGTCTCT	247					
954	TCTCATTTATAGGGACTGGAAATGCMAAAGTAGGAAGCAAGAAACACCTGGAGTAACAGG	1013					
246	TTCATTATAGGCACATGGAATGCMAAAGTAGGAAGTCAAGAAACACCTGGAGTAACAGG	187					
1014	CAAAATTTGGCTTGGAAATACGGAATGAAGCAGGGCAAGACTAATAGAGTTTGGCCAGA	1073					
186	CAAAATTTGGCTTGGAAATACGGAATGAAGCAGGGCAAGACTAATAGAGTTTGGCCAGA	127					
1074	AAATGCATCTGGTCAATGACAAACACCTCTTCCAAACCAACCAAGAGAAAGACTCTACACATG	1133					
126	AAATGCATCTGGTCAATGACAAACACCTCTTCCAAACCAACCAAGAGAAAGACTCTACACATG	67					
1134	GACATCACAGATGGTCAACCGAAATCAGATGATATATTCTTTGGAGCCAAA	1189					
66	GACATCAACAGATGGTCAACCGAAATCAGATGATATATTCTTTGGAGCCAAA	11					

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

1 (Bases 1 to 947)

Holt, R., Stott, J., Yang, G., Barber, S., Smailus, D., Prabhu, A.-L., Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M., Butlerfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chin, R., Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S., Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W., Dalrymple, B. P. and Tellam, R.

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398 Unpublished (2003)

Other GSSs: CH240_392111.TARBAC13P2

Contact: Rob Holt

Sequencing

The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6

Tel: 604-877-6085
Fax: 604-877-6276

Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240 (<http://www.chori.org/bacpac/bovine240.htm>). For BAC library availability, please contact Peter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Science Centre, Canada.

FEATURES

```

taxon="taxon:9913"
/ds_xref="taxon:9913"
/clone="CH240_392f11"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note=vvector: vPARBAC1.3; Site.1: MboI; Site.2: MboI;
Hereford bull Ll Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match      21.5%; Score 769.501; DB 1; Length 947;
Best Local Similarity 91.2%; Pred.No. 0.18;
Matches 869; Conservative 0; Mismatches 75; Indels 9; Gaps 5;

```

RESULT 37	CC591633	947 bp	DNA	linear	GSS 18-JUN-2003
CC591633/c	CH240_392111.T7	CHORI-240	Bos taurus	genomic clone	CH240_392111,
LOCUS	CH240_392111				
DEFINITION	genomic survey sequence.				
ACCESSION	CC591633				
VERSION	CC591633.1	GI:31949515			
KEYWORDS	GSS.				
SOURCE	Bos taurus	(cow)			
ORGANISM	Bos taurus				


```

Db      661  TGTCTTTTCATTATAGGGGACTGGAATGCTAAAGTAGGAAGTCAAGAACTCTCTGGAGT 720
QY      1008  AACAGCAAAATTTGGCTTTGGAATACGGAATGAAGCAGCGGCAAGACTTAATAGAGTTTGG 1067
Db      721  AACAGCAAAATTTGGCTTTGGAATACGGAATGAATAGGCGCAAGACTTAATACAGTTTGG 780
QY      1068  CCAAGAAATGCACTGGTCATAGCAAAACACCTCTTCCAAACAACACACAGAGAGACTCTTA 1127
Db      781  -CAAGAAATACACTGATCATTAACAACAACCTCTTCCAAACAACAAGAGAGACTCTTA 839
QY      1128  CACATGGACATCACCAGAT 1146
Db      840  CACATGGACATCACCAGAT 858

CC525458      877 bp      DNA      linear      GSS 17-JUN-2003
LOCUS      CH240_399H6.T7 CHORI-240 Bos taurus genomic clone CH240_399H6,
DEFINITION      genomic survey sequence.
ACCESSION      CC525458
VERSION      CC525458.1 GI:31843746
KEYWORDS      GSS.
SOURCE      Bos taurus (cow)
ORGANISM      Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 877)
Holt, R., Scott, J., Yang, G., Barber, S., Smalhus, D., Prabhu, A.-L.,
Tsai, M., Cloutier, A., Lee, D., Girm, N., Olson, T., Mayo, M., Chiu, R.,
Butterfield, J., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
Schein, J., Marra, M., de Jong, P., Keele, J.W. and Kappes, S.M.
Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
Unpublished (2003)
Other GSSs: CH240_399H6.TARBAC13P2
Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
Tel: 604-877-6085
Fax: 604-877-6276
Email: rholt@bcgsc.ca
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@gmail.com).
Clones may be purchased from BACpac Resources
(http://www.chori.org/bacpac/ordering_information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.
Plate: 399 row: H column: 6
Seq primer: T7
Class: BAC ends.
FEATURES
    source
        1..877
            /organism="Bos taurus"
            /mol_type="genomic DNA"
            /strain="Breed: Hereford"
            /db_xref="taxon:9913"
            /clone="CH240_399H6"
            /sex="Male"
            /cell_type="Blood"
            /clone_lib="CHORI-240"
            /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
            Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
            library (Male) produced by Pieter de Jong"
Query Match      21.5%; Score 789.5; DB 1; Length 877;
Best Local Similarity 94.0%; Fred. No. 0.2;
Matches 839; Conservative 0; Mismatches 46; Indels 7; Gaps 3;

```

```

QY      344  TTGAGAACCCCATGAACAGTATGAAAGGC-AAAATGATAGGATCTGAAAGAGAACTC 402
Db      877  TTGAGAACCCCATGAACAGTATGAAAGGC-AAAATGATAGGATCTGAAAGAGAACTC 818
QY      403  CCAGGTCACTAGTGGTCCCATATGCTCTAGTGGATCAGTGGAGAAATACTCCAGAAAG 462
Db      817  CCAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 758
QY      463  AATGAAGAGATGGAGCCCAAGCAAAAGCAATATCCAGCTGTGGATGTGACTGGTATATA 522
Db      757  AATGAAGGATGGAGCCCAAGCAAAAGCAATATCCAGCTGTGGATGTGACTGGTATATA 698
QY      523  AGCAAGGTCGATGCTGTAAAGAGCAATATGCAATAGGAACCTGGAATGTCAGTCCATG 582
Db      697  AGCAAGGTCGATGCTGTAAAGAGCAATATGCAATAGGAACCTGGAATGTCAGTCCATG 638
QY      583  AATCAAGGCAAAATTCGAATGGTCAACAAGAGATGGCAAGAGTGAATGTCAACATTTCTA 642
Db      637  AATCAAGGCAAAATTCGAATGGTCAACAAGAGATGGCAAGAGTGAATGTCAACATTTCTA 578
QY      643  GGAATCAGCGAACTAAATGGATGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 702
Db      577  GGAATCAGCGAACTAAATGGATGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 518
QY      703  TACTACTGGGCGAGAAATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAGA 762
Db      517  TACTACTGGGCGAGAAATCCCTCAGAAAGAAATGGAGTAGCCATCATGGTCAACAAAGA 458
QY      763  GTCCCAATGAGTACTTGGATGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 822
Db      457  GTCTGAATGAGTACTTGGATGGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 398
QY      823  CAAGCAAAACCAATTCATATCACAGTAAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAA 882
Db      397  CAAGCAAAACCAATTCATATCACAGTAAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAA 338
QY      883  GAAGTGAAGTGAACGCTCTATGAAGACCTTACAGACCTTTTAGACCTAACACCCAAA 942
Db      337  GAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 278
QY      943  AAAGATGTCCTTCTCATTTATAGGGGACTGGAAATCAAAAGTAGGAAGCAAGAAACACCT 1002
Db      277  AAAGATGTCCTTCTCATTTATAGGGGACTGGAAATCAAAAGTAGGAAGTGAAGTGAAGTGAAG 218
QY      1003  GGAGTGAACGCAAAATTTGGCTTGGATAGCGATAGCGATAGCGATAGCGATAGCGATAGCGAT 1062
Db      217  GGAGTGAACGCAAAATTTGGCTTGGATAGCGATAGCGATAGCGATAGCGATAGCGATAGCGAT 158
QY      1063  TTTTCCCAAGAAATGCACTGGTCAATAGCAAAACACCTCTTCCCAACCAAGAGAGAAGA 1122
Db      157  TTTTCCCAAGAAATGCACTGGTCAATAGCAAAACACCTCTTCCCAACCAAGAGAGAAGA 103
QY      1123  CTCTACATGAGACATACC-AGATGTCACACCGAAATCAGATTGATTATTTCTTTG 1181
Db      102  TTCTACATGAGATATCACCAGATGGTCAACACCGAAATCAGATTGATTATTTCTTTG 43
QY      1182  CAGCCAAAGATGGAGAGCTCTATACAGTCAAGCAAAACCAAG 1223
Db      42  CAGCCAAAGATGGAGATGCTCTATACAGTCTACAAACCAAG 1

RESULT 40
CC548709/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

```

CC548709 841 bp DNA linear GSS 17-JUN-2003
 CH240_432B6.T7 CHORI-240 Bos taurus genomic clone CH240_432B6,
 genomic survey sequence.
 CC548709
 CC548709.1 GI:31866993
 GSS.
 Bos taurus (cow)
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;


```

FEATURES
  source
    Class: BAC ends.
    Location/Qualifiers
      1..865
        /organism="Bos taurus"
        /mol_type="genomic DNA"
        /strain="breed: Hereford"
        /db_xref="taxon:9913"
        /clone="CH240_412K21"
        /sex="Male"
        /cell_type="Blood"
        /clone_lib="CHORI-240"
        /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
        Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
        library (Male) produced by Pieter de Jong"

Query Match      21.5%; Score 768.4; DB 1; Length 865;
Best Local Similarity 94.5%; Pred. No. 0.2;
Matches 817; Conservative 0; Mismatches 46; Indels 2; Gaps 2;

QY 202 CACTAGGACACAGCCTTGTCTTAATCAATCAATGAAA-CTAAGCCATGCCCGTGGGGCAACCC 260
DB 865 CACTAGCACCACAGCCTTGTCTTAATCAATCAATGAAAACCTAAGCCATGCCCGTGGGGCAACCC 806
QY 261 AGATGGGAGCTCATGTGGAGAGATCTGACAGAAATGTGTCTCACTGGAGAGGGAATG 320
DB 805 AGACAGGGGGGTCTGTGGTGGAGAAATCTGACAGAAATGTGTCTCACTGGAGAGGGAATG 746
QY 321 -CAAAACCACTTCAGTATTTCTTGCCTTTGAGAACCCCATGACAGTATGAAAAGGCCAAATG 379
DB 745 TCARACCACTTCAGTATTTCTTGCCTTTGAGAACCCCATGACAGTATGAAAAGGTTAAATG 686
QY 380 ATAGTACTGAAAGAGNACTCCCGAGTCTAGTAGTGCCCATATGCTACTGGAGATC 439
DB 685 ATAGATACTGAAAGAGNACTCCCGAGTCTAGTAGTGCCCATATGCTACTGGAGATC 626
QY 440 AGTGAGAAATTAATCTCCAGAAAGATGAAGAGATGGAGCCAAAGCAAAAGAAATACCCAG 499
DB 625 AGTGAGAAATTAATCTCCAGAAAGATGAAGAGATGGAGCCAAAGCAAAAGAAATACCCAG 566
QY 500 CTGTGATGTGACTGGTGATTAACAGAGGTCCGATGCTGTAAAGAGCAATATTGATAG 559
DB 565 CTGTGATGTGACTGGTGATTAACAGAGGTCCGATGCTGTAAAGAGCAATATTGATAG 506
QY 560 GAACCTGGAATCTCAGTGCATTAATCAAGGCAATTCGAAGTGGTCAACAGAGATGG 619
DB 505 GAACCTGGAATCTCAGTGCATTAATCAAGGCAATTCGAAGTGGTCAACAGAGATGG 446
QY 620 CAAGATGATGCTCAATCTTAGGAATCAGCGAACTAAATGGAATGGAGTGGATGGAT 679
DB 445 CAAGATGGAACACTGACATCTTAGGAATCAGCGAACTAAATGGAATGGAGTGGATGGAT 386
QY 680 TTAACTCAGATCACCATTATATCTACTCTGGGGCAGCAATCCCTCAGAGAAATGGAG 739
DB 385 TTAACTCAGATCACCATTATATCTACTCTGGGGTAGATTCCTCAGAGAAATGGAG 326
QY 740 TAGCCATATGTCAACAAAGAGTCCGAAATGCAATCTTGGATGCAATCTCAAAACG 799
DB 325 TAGCCATATGTTCACAAAGAGTCCAAATGCAATCTTGGATGCAATCTCAAAACG 266
QY 800 ACAGATGATCTCTGTTTGTTCCTCAAGGCAACCAATCAATATCAGATGATCAAGTCT 859
DB 265 ACAGATGATCTCTGTTTGTTCCTCAAGGCAACCAATCAATATCAGATGATCAAGTCT 206
QY 860 ATGCCCAACACAGTAAATCTGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACCTCAAG 919
DB 205 ATGTCACCAACAGTAAATCTGAAGAGCTGAAGTTGAACGGTCTCTATGAAGACCTCAAG 146
QY 920 ACCTTTTGAACCTAACACCCAAAGAGTCTCTTCTCATTTATAGGGGACTGGGAATCAA 979
DB 145 ACCTTTTGAACCTAACACCAAGAGATGTCCTTTCTCATTTATAGGGGACTGGGAATCAA 86
QY 980 AAGTAGGAAGCAAGAAACACCTGAGTAACAGGCAAAATTTGGCTTGGATACGGAATG 1039

```

```

DB 85 AAGTAGGAAGTCAAGAAACACCTGGAGTAACAGGCAAAATTTGGCCTTAGAATACAGAATG 26
QY 1040 AAGCAGGCAAGAGACTAATAGATT 1064
DB 25 AAGCAGGCAAGAGACTAATAGATT 1

RESULT 42
CC581625/c
LOCUS
DEFINITION
  CH240_376F14.T7 CHORI-240 Bos taurus genomic clone CH240_376F14,
  genomic survey sequence.
ACCESSION
  CC581625
VERSION
  CC581625.1 GI:31929477
KEYWORDS
  GSS.
SOURCE
  Bos taurus (cow)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
    Bovidae; Bovinae; Bos.
  REFERENCE
    1 (bases 1 to 822)
    Holt, R., Stott, J., Yang, G., Barber, S., Smalley, D., Prabhu, A.-L.,
    Tsai, M., Cloutier, A., Lee, D., Ginn, N., Olson, T., Mayo, M., Chiu, R.,
    Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
    Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
    Schein, J., Marra, M., de Jong, P., McWilliam, S., Barris, W.,
    Dalrymple, B.P. and Tellam, R.
    Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
    Unpublished (2003)
  JOURNAL
    Other GSSs: CH240_376F14.TARBAC13P2
  COMMENT
    Contact: Rob Holt
    Sequencing
    The British Columbia Cancer Agency Genome Science Centre
    600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4S6
    Tel: 604-877-6085
    Fax: 604-877-6276
    Email: rholt@bcsc.ca
    Clones are derived from the bovine BAC library CHORI-240
    (http://www.chori.org/bacpac/bovine240.html). For BAC library
    availability, please contact Pieter de Jong (pjejong@email.cho.org).
    Clones may be purchased from BACPAC Resources
    (http://www.chori.org/bacpac/ordering_information.htm). This work
    was undertaken as part of the International Bovine BAC Mapping
    Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
    British Columbia Genome Sciences Centre, Canada.
    Plate: 376 Row: F Column: 14
    Seq primer: T7
    Class: BAC ends.
    Location/Qualifiers
      1..822
        /organism="Bos taurus"
        /mol_type="genomic DNA"
        /strain="breed: Hereford"
        /db_xref="taxon:9913"
        /clone="CH240_376F14"
        /sex="Male"
        /cell_type="Blood"
        /clone_lib="CHORI-240"
        /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
        Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
        library (Male) produced by Pieter de Jong"

Query Match      21.5%; Score 767.3; DB 1; Length 822;
Best Local Similarity 96.6%; Pred. No. 0.21;
Matches 794; Conservative 0; Mismatches 27; Indels 1; Gaps 1;

QY 408 GTCAGTAGTGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGATGA 467
DB 822 GTCAGTAGTGCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCAGAAAGATGA 763
QY 468 AGAGATGGAGCCAAAGCAAAAGAAATACCCAGCTGTGGATGTGCTGATATACGAA 527
DB 762 AGGATGGAGCCAAAGCAAAAGCAAAATACCCAGCTGTGGATGTGCTGATATACGAA 703

```

QY	528	GGTCCGATGCTCTAAAGAGCAATATTGCGATAGGAACCTGGAAATGTCAGGTCCATGAATCA	587
Db	702	GGTCCGATGCTGTAAGAGCAATATTGCGATAGGAACCTAGAAATGTCAGGTCCATGAATCA	643
QY	588	AGGCCAAATTTGGAAATGGTCAAAACAAGAGATGCCAAGAGTGAATGTCAACATTCTAGGAAT	647
Db	642	AGGCCAAATTTGGAAATGGTCAAAACAAGAGATGCCAAGAGTGAATGTGCGACATTCTAGGAAT	583
QY	648	CAGCGAACTAAAAATGCAATCTGGAATGGGTGAATTTTAATTCAGATGACCACTTATATCTACTA	707
Db	582	CAGCGAACTGAAAATGGAATGGGTGAATTTTAATTCAGATGACCACTTATATCTACTA	523
QY	708	CTGGGGCAGGAATCCCTCAGAGAATGGAGTAGCCATCATGTGTCAACAAAGAAGTCCG	767
Db	522	CTGTGGGCAGGAATCCCTCAGAGAATGGAGTAGCCATCATGTGTAAACAAGAGATCCG	463
QY	768	AAATGCGAGTACTTGATGCACTCTCAAAAACGACAGAGATGATCTCTGTTCTTTTCCCAAGG	827
Db	462	AAATGCGAGTACTTGATGCAATCTCAAAAACGACAGAGATGATCTCTGTTCTTTTCCCAAGG	403
QY	828	CAAAACCAATTCAAATATCAGATATCCAAAGTCTATGCCCCAAACCAAGTAAATGCTGAAGAAGC	887
Db	402	CAAAACCAATTCAAATATCAGATATCCAAAGTCTATGCCCCAAACCAAGTAAATGCTGAAGAAGC	343
QY	888	TGAAGTTGAACCGGTCTTATGAAGACCTTACAAGACCTTTTGAACCTAACACCCCAAAAAGA	947
Db	342	TGAAGTTGAACCGGTCTTATGAAGACCTTACAAGACCTTTTGAACCTAACACCCCAAAAAGA	283
QY	948	TGTCCTTCTCATTTATAGGGAGCTGGAATGCAAAAGTAGGAAGCAAGAAACACCTGGAGT	1007
Db	282	TGTCCTTCTCATTTATAGGGAGCTGGAATGCAAAAGTAGGAAGCTAATAGAGTTTTC	223
QY	1008	AACAGGCAAAATTTGGCCTTGGAATACCGAATGAAGCAGGCGCAAGACTAATAGAGTTTTC	1067
Db	222	AACAGGCAAAATTTGGCCTTGGAATACCGAATGAAGCAGGCGCAAGACTAATAGAGTTTTC	163
QY	1068	CCAAAGAAATGCACTGTGCTATAGCAAAACACCTCTTCCAAACAACACAAAGAGAAGACTCTA	1127
Db	162	CCAAAGAAATGCACTGTGCTATAGCAAAACACCTCTTCCAAACAACACAAAGAGAAGACTCTA	103
QY	1128	CACATGACATCACCAGATGGTCAACACCCGAATTCAGATTGATTATATCTTTGCGAGCC	1186
Db	102	TACATGACATCACCAGATGTTCAACACCCGAATTCAGATTGATTATATCTTTGCGAGCC	43
QY	1187	AAAGATGGAGAAGCTCTATATCAGTCTAGCAAAAAACAAGACCAG	1228
Db	42	AAAGATGGAGAAGCTCTATATCAGTCTAGCAAAAAACAAGTCCAG	1

Contact: Rob Holt
Sequencing
The British Columbia Cancer Agency Genome Science Centre
600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
Tel: 604-877-6085
Fax: 604-877-8276
Email: rholt@bcgsc.ca

Clones are derived from the bovine BAC library CHORI-240
(<http://www.chori.org/bacpac/bovine240.htm>). For BAC library
availability, please contact Pieter de Jong (pdjong@email.cho.org).
Clones may be purchased from BACPAC Resources
(<http://www.chori.org/bacpac/orderinginformation.htm>). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBMC) by CSIRO Livestock Industries, Australia and the
British Columbia Genome Sciences Centre, Canada.

Plate: 396 row: L column: 24
Seq primer: SP6
Class: BAC ends.

Location/Qualifiers
1. 861
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_396L24"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: Mbol; Site 2: Mbol;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"

Query Match	21.5%	Score 767	DB 1	Length 861
Best Local Similarity	94.5%	Pred. No. 0.2		
Matches 815	Conservative 0	Mismatches 45	Indels 2	Gaps 2
QY	440	AGTGGAGAAATACTCCAGAAAGAAATGAAGAGATGGAGCCAAAGCAAAAAGAATATCCCG 499		
DB	1	AGTGGAGAAATACTCCAGAAAGAAATGATTGGATGGAGCCAAACAAAAACATATCCCG 60		
QY	500	CTGTGGATGTGACTGTGGTGATATAAGCAAGGTCCGATGCTGTAAAGAGCAATATTGCATAG 559		
DB	61	TTGTGGATGTGACTGGTGATAGAAGCAGGTCCGATGCTGTAAAGAGCAATATTGCATAG 120		
QY	560	GAACCTGGAAATGTCAGTCCATGAATCAAGGCAAAATTGGAAGTGGTCAACAAAGATGG 619		
DB	121	GAACCTGGAAATGTTAGTCCATGAATCAAGGCAAAATTGAAAGTAGTCAAACTGGAGATGG 180		
QY	620	CAAGAGTGAATGCAACATTTCTAGGAATCAGCGAACTTAAATCGCATGCAATGGGTGAAT 579		
DB	181	CAAGAGTGAATGTGCACATTTCTAGAAATCAGCNAATCTAAATGGATGAAATGGGTGAAT 240		
QY	680	TTAACTCAGATGACCAATTATATCTACTCTGGGGGAGGAATCCCTCAGAAGAAATGGAG 739		
DB	241	TTAACTCAGATGACCAATTATATCTACTCTGTGGGAGGAATCCCTTAGAAGAAATGGAG 300		
QY	740	TAGCCATCATGTTCAACAAAGAGTCGAAATTCGAGTACTTGGATGCAGTCTCAAAAAG 799		
DB	301	TAGCCATCAT - GTCAACAAAGAGTCTGAAATTCGAGTACTTGGCTGCAATCTCAAAAAGA 359		
QY	800	ACAGAATGATCTCTGTTCCTTTCCAAAGGCAACCAATTCAAATATCACAGTAATCCCAAGTCT 859		
DB	360	ACAGAATATCTCTGTTCTTTCCAAAGGCAACCAATTCATAGACAGTAATCCCAAGTCT 419		
QY	860	ATGCCCCAAACAGTAATGCTGAAGAAGCTGAAAGTTGAAACGGTCTTATGAAGACCTTACAAG 919		
DB	420	ATGCTCCAAACAGTAATGCTGAAGAAGCTGAAAGTTGAAACAGTCTTATGAAGACCTTACAAG 479		
QY	920	ACCTTTTAGAACTTAACACCCAAAAAGATGTCCTTCTCATTTATAGGGGACCTGGAAATGCAA 979		
DB	480	ACCTTTTAGAACTTAACACCCAAAAAGATGTCCTTTTCATTTATAGGGGACCTGGAAATGCAA 539		
QY	980	AAAGTAGGAACGAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCCCTTGGAAATACGGAATG 1039		

```

Db      540  AAGTAGGAAGTCAAGAAACCTGGGGTAACAGGCAAAATTTGGGCTTGGAAATGGAATG 599
QY      1040  AGCAGGGCAAGACTAATAGAGTTTTCAGAGAAATGCACTGGTCATAGCAACACCC 1099
Db      600  AAGCAGGGCAAGGGCTAATAGAGTTTTCAGAGAAATGCACTGGTCATAGCAACACCC 659
QY      1100  TCTTCCAAACACACAGAGAGACTCTACATGAGACATCACCAGATGGTCACACCGAA 1159
Db      660  TCTTCCAAACACACAGAGAGACTCTACATGAGACATCACCAGATGGTCACACCGAA 719
QY      1160  ATCAGATTGATTATATCTTTTCAGCCAAAGATGGAGAGCTCTATACAGTCAGCAAAAA 1219
Db      720  ATCAGATTGATTATATCTTTTCAGCCAAAGATGGAGAGCTCTATACAGTCAGCAAAAA 779
QY      1220  CAAGACC-AGGAGCTTACTGTGGCTCAGATCATGAACTCTTATTCGCAATTCAGACTT 1278
Db      780  CAAGCCCGTGGAGCTGACTGTGGCTCAGATCATGAACTCTTATTCGCAATTCAGACTT 839
QY      1279  AAATTGAAGAAAGTAGGGA AAA 1300
Db      840  AAATTGAAGAAAGTAGGGA AAA 861

```

RESULT 44

CC512505

LOCUS

DEFINITION

CH240_356L21.TARBAC13P2 CHORI-240 Bos taurus linear GSS 17-JUN-2003

CH240_356L21, genomic survey sequence.

CC512505

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CC512505 870 bp DNA linear GSS 17-JUN-2003
 CH240_356L21.TARBAC13P2 CHORI-240 Bos taurus genomic clone
 CH240_356L21, genomic survey sequence.
 CC512505.1 GI:31830793
 GSS-
 Bos taurus (cow)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 1. (bases 1 to 870)
 Holt, R., Stott, J., Yang, G., Barber, S., Smalhus, D., Prabhu, A.-L.,
 Tsai, M., Cloutier, A., Lee, D., Ginn, N., Olson, R., Mayo, M.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Maira, M., de Jong, P., McWilliam, S., Barris, W.,
 Dalrymple, B.P. and Tellam, R.
 Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
 Unpublished (2003)
 Other_GSSs: CH240_356L21.T7
 Contact: Rob Holt
 Sequencing
 The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4B6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bcgsc.ca
 Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pdejong@email.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering/information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 356 row: L column: 21
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 1. 870
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="bred: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_356L21"
 /sex="Male"

FEATURES

source

1. 870
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="bred: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_356L21"
 /sex="Male"

RESULT 45
 CCS40150

```

/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"
Query Match      21.5%; Score 766.4; DB 1; Length 870;
Best Local Similarity 93.9%; Pred. No. 0.2;
Matches 818; Conservative 0; Mismatches 51; Indels 2; Gaps 2;
QY      289  TGACAGAATGTGGTCCACTGGAGAAGGAAT-GCAAAACCACTTCAGTATTCTTGCCTTGA 347
Db      1    TGACAGAAGGTGGTCCGCTGGAGAAGGAATGGCAATCCACTTCAGTATTCTTGCCTTGA 60
QY      348  GBAACCCCATGACAGTATGAAAGCGCAAAATGATAGGATCTGAAAGAGAACTCCCCAG 407
Db      61  AAACCCCATGACAGTATCAAAAGCGCAAAATGATAGGATCTGAAAGGGGAACCTCCCG 120
QY      408  GTCAGTAGGTGCCCCATATGCTACTGGAGATCAGTGGAGAAATAACTCCCAAGAAATGA 467
Db      121  GTCAGTAGGGACCCCAATATGCTACTGGATATCAGTGGAGAAATAACTCCAGAAAGATGA 180
QY      468  AGAGATGGAGCCAAAGCAAAAGAAATACCCAGCTGTGGATGTGACTGGTGTATATAGCAA 527
Db      181  AGGATGGAGCCAAAGCAAAACCAATACCCAGTTCGAGTTGTGACTGGTGCACAGAAACAA 240
QY      528  GGTCCGATGCTGTAAAGAGCAATATTGCTAGGAACCTGGAATGTCAGGTCCATGAATCA 587
Db      241  GGTCTGTGCTGTAAAGAGCAATATTGCTAGGAACCTGGAATGTCAGGTCCATGAATCA 300
QY      588  AGGCAAAATGGAAGTGTCAAAACAGAGATGGCAAGAGTGAATGTCAAATTTCTAGGAAT 647
Db      301  AGGCAAAATGGAAGGGGTCAAAACAGAGATGGCAAGAGTGAATGTCAGCAATTTCTAGGAAT 360
QY      648  CAGCGAACTAAATAGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTA 707
Db      361  CAGCGAACTAAATAGGACTGGAATGGGTGAATTTAACTCAGATGACCATTTATATCTACTA 420
QY      708  CTGCGGGCAGGAATCCCTCAGAGAAATGAGTAGCCATCATGTCGTCACAAAAGAGTCCG 767
Db      421  CTGCGAGGAGGAATCCCGCAGAGAAATGGAAGTAGCCATCATGTCGTCACAAAAGAGTCCG 480
QY      768  AAATGCACTACTTGGATGCACTCTCAAAACGACAGAAATGATCTCTCTTTGTTTCCAAAG 827
Db      481  AAATGCACTACTTGGATGCACTCTCAAAACGACAGAAATGATCTCTCTTTGTTTCCAAAG 540
QY      828  CAAACCATTTCAATATCAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAGAGC 887
Db      541  CAAACCATTTCAATATCAGTAATCCAAAGTCTATGCCCAACCAAGTAATGCTGAAGAGC 600
QY      888  TCAAGTTCAACCGTCTTATGAGACCTTACAAAGACCTTTTAGAACTAACACCCCAAAAGA 947
Db      601  TGAAGATGAACGGTTCTATGAGACCTTACAAAGTCTTATGGAAT-ATACCCCAAAAGG 659
QY      948  TGTCTTTCTCATTTATAGGGACTGGAAATGCAAAAGTAGGAAGCAAAAGAACACCTGGAGT 1007
Db      660  TGTCTTTCTCATTTATAGGGACTGGAAATGCAAAAGTAGGAAGTACAGAAACACCTGCAAT 719
QY      1008  AACAGGCAAAATTTGGCTTGAATACGGAATGAAGCAGGGCAAGAGCTAATAGAGTTTTC 1067
Db      720  AACAGGCAAAATTTGGCTTGAATACGGAATGAAGCAGGGCAAGAGCTAATAGAGTTTTC 779
QY      1068  CAAAGAAATGCACTGGTCTATAGCAAAACACCTCTTCCAAACACAGAGAGAGACTCTA 1127
Db      780  CAAAGAAATGCACTGGTCTATAGCAAAAGACCTTTTCCAAACACAGAGAGAGACTCTA 839
QY      1128  CACATGGACATCACCAGATGGTCAACACCGA 1158
Db      840  CACATGGACATCACCATATGGTCAACACCGA 870

```

LOCUS CC540150 862 bp DNA linear GSS 17-JUN-2003
 DEFINITION CH240_420F7.TARBAC13P2 CHORI-240 Bos taurus genomic clone
 CH240_420F7, genomic survey sequence.
 ACCESSION CC540150
 VERSION CC540150.1 GI:31858434
 KEYWORDS GSS.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 862)
 AUTHORS Holt, R., Stott, J., Yang, G., Barber, S., Smalus, D., Prabhu, A.-L.,
 Tsai, R., Cloutier, A., Lee, D., Girn, N., Olson, T., Mayo, M.,
 Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,
 Mathewson, C., Wye, N., Masson, A., Brown-John, M., Jones, S.,
 Schein, J., Marra, M., de Jong, P., Keele, J. W. and Kappes, S. M.
 TITLE Bovine BAC End Sequences from Library CHORI-240, PLATES 399 to 478
 JOURNAL Unpublished (2003)
 COMMENT Other GSSs: CH240_420F7.T7
 Contact: Rob Holt
 Sequencing The British Columbia Cancer Agency Genome Science Centre
 600 W. 10th Ave. Vancouver, British Columbia, Canada V5Z 4E6
 Tel: 604-877-6085
 Fax: 604-877-6276
 Email: rholt@bcgsc.ca
 Clones are derived from the bovine BAC library CHORI-240
 (<http://www.chori.org/bacpac/bovine240.htm>). For BAC library
 availability, please contact Pieter de Jong (pdejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering_information.htm). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBMC) by CSIRO Livestock Industries, Australia and the
 British Columbia Genome Sciences Centre, Canada.
 Plate: 420 row: F column: 7
 Seq primer: SP6
 Class: BAC ends.

FEATURES
 source Location/Qualifiers
 1..862
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="breed: Hereford"
 /db_xref="taxon:9913"
 /clone="CH240_420F7"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"
 /note="vector: PTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull Li Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

Query Match 21.4%; Score 765.9; DB 1; Length 862;
 Best Local Similarity 94.3%; Pred. No. 0.2;
 Matches 815; Conservative 0; Mismatches 46; Indels 3; Gaps 2;

QY 320 GCAACACATTCAGTATTCTTGCTTGAGAACCCCATGACAGTATGAAAGGCAAAATG 379
 |||||
 DB 1 GCAATCATTGATTTCTTGCTTGAGCACCCTGACAGTATGAAAGGCAAAATG 60
 |||||
 QY 380 ATAGGATCTGAAAGAGGAAGT-CCCCAGTCTAGTGGTCCCCCATATGCTACTGGAGAT 438
 |||||
 DB 61 ATAGGATCTGAAAGAGGAAGTCCCCAGTCTAGTGGTCCCCCATATGCTACTGGAGAT 120
 |||||
 QY 439 CAGTGGAGAAATACCTCCGAAAGATGAGAGATCGAGCCCAAGCAAAAGAAATACCCA 498
 |||||
 DB 121 CAGTGGAGAAATACCTCCGAAAGATGAGAGATGCTGCTGAAAGAGCAATATTCGATA 180
 |||||
 QY 499 GCTGTGGATGTGCTGCTGATATTAAGCAAGGTCGATGCTGTAAGAGAGCAATATTCGATA 558
 |||||
 DB 181 GCTGTGGATGTGCTGCTGATATTAAGCAAGGTCGATGCTGTAAGAGAGCAATATTCGATA 240
 |||||
 QY 559 GGAACCTGGAAATGTCAGGTCCATGATCAAGGCAAAATGGAAGTGGTCAACAGAGATG 618
 |||||

DB 241 GGAACCTGGAAATGTTAGGTCCATGATCAAGGCAAAATGGAAGTGGTCAACAGAGATG 300
 |||||
 QY 619 GCAAGAGTGAATGTCAACATTTCTAGGAATCAGCGAACTAAAAATGGCATGGGTGAA 678
 |||||
 DB 301 GCAAGAGTGAATGTGACATTTCTAGGAATCAGCGAACTAAAAATGGCATGGGTGAA 360
 |||||
 QY 679 TTAACTCAGATGACCACTTATATCTACTTCTGGGGCAGGAATCCCTCAGAAGAAATGGA 738
 |||||
 DB 361 TTAACTCAGATGACCACTTATATCTACTTCTGGGGCAGGAATCCCTCAGAAGAAATGGA 420
 |||||
 QY 739 GTAGCCATCATGGTCAACAAAAAGAGTCCGAAATGCAGTACTTGGATGCAGTCTCAAAAAC 798
 |||||
 DB 421 GTGGCCATCATGGTCAACAAAAAGAGTGTAGAAATGCAGTACTTGGATGCAGTCTCAAAAAT 480
 |||||
 QY 799 GACAGATGATCTCTGTTGTTTCCAAAGGCAAAACCAATTCATATACAGTATCCAAATC 858
 |||||
 DB 481 GACAGATGATCTCTGTTGTTTCCAAAGGCAAAACCAATTCATATACAGTATCCAAATC 540
 |||||
 QY 859 TATGCCCCCAACCAAGTAAATGCTGAAGAACTGAAGTTGAACGCTCTATGAAGACCTACAA 918
 |||||
 DB 541 TATGCCCCCAACCAAGTAAATGCTGAAGAACTGAAGTTGAATGGTCTATGAAGAACTACAA 600
 |||||
 QY 919 GACCTTTTAGAACTAACACCCCAAAAAAGATGTCCTTCTCATATAGGGGACTGGAAATGCA 978
 |||||
 DB 601 GACCTTTTAGAACTAACACCCCAAAAAAG--GTCTCTTTTCATTATAGGTGACTGGAATGCA 658
 |||||
 QY 979 AAAGTAGGAAGCAAGAAACACCTGGAGTAAACAGGCAAAATTTGGCTTGGATACGGAT 1038
 |||||
 DB 659 AAAGTAGGAAGTCAAGAAACACCTGGAAATAACAGGCAAAATTTGGCTTGGAGTACAGACT 718
 |||||
 QY 1039 GAAGCAGGGCAAGAACTAATAGAGTTTTCGAAGAAATGCACTGGTTCATAGCAAAACACC 1098
 |||||
 DB 719 GAAGCAGGGCAAGAGCTAATAGAGTTTTCGAAGAAAGACACTGGTTCATAGCAAAACACC 778
 |||||
 QY 1099 CTCCTTCCAAACACACAGAGAAAGTCTACACATGGAGCATCCAGATGGTTCACACCGA 1158
 |||||
 DB 779 CTCCTTCCAAACACACAGAGAAAGTCTTATACATGGAGCATCCCTGATGGTTCACACCGA 838
 |||||
 QY 1159 AATCAGATTGATTATATTCTTTGC 1182
 |||||
 DB 839 AAACAGATGGATTATATTCTTTGC 862
 |||||

Search completed: August 24, 2004, 13:41:22
 Job time : 79 secs


```
QY 530 AGTCTTGTGTTTG-GTGAATAGTCTGTAATATCT-CTAGTCCACTTGGTTATGA 587
Db 53696 ATTCTTTTGTCTTGTAGATTGAATGTTCTATAAATATCTGTTAGATCCATTGGTTTCATAG 53637
QY 588 CATCAGTTAGCTCCAGCATTTCTGTTTTCGTTTGTGTTGAGATGACCTAACTGTTGGA 647
Db 53636 CATCTGTTACTCTTCACTGCTGCTCTGTTTAG-TTTGTTGTTTCCAAGATCTGTCATTTGCT 53578
QY 648 GAGATGGGGTATTGAAGTAGCCCACTATCTGCTGTGAGGT-CAATATGATGATTTTACG 706
Db 53577 GAGAGTAGGGTGTGGAAGTCTCCCACTATTATTGTTGTTGCTGCAAGGTGCTTTTCAGC 53518
QY 707 TGTAGCTGCTGTTGTTTATGAACCTTGGTGACATTTGTTGTTGTTGATGACATTAAGA 766
Db 53517 TTTAGTAAAGTTTCTTTTATGAATGTTGGTGCCTTGCATTGGAGTATAGATGTTGAGA 53458
QY 767 ATTGCAATGTCCTCTTGTGTGA-TTTTCTTGTGATGCTATGTAGTATTTCTTCCCAATCT 825
Db 53457 ATTTAGAATGCATCTTGTGTAGATTTTCTTTGATGAGAATGAAGTATCTTCTCTTATTC 53398
QY 826 CA---TCTGCTTAGTTTGGGTTTAACTCTATTAGTCAAGATTTAAATGACTGTAT--- 879
Db 53397 TATTTTGTAAACCTTTGGTTGAAGTTGATTTTATTGATAGTAAATGAATGATCT 53338
QY 880 -CGGCTTCTCTTAGGGCCATTGCTTAGAATA-TCTTTTCCATCTCTTTTACTCTAAGG 937
Db 53337 CCAGCTGTTTCTTGGGACCAATTTGCTTTGGAAATTTGTTTCAAGCCCTTTACTCTGAGG 53278
QY 938 -----TGATGTCATACCATGGTAGTGTCTTTTGGATGCGACGAGTAGGATGATCT 991
Db 53277 TAGTACTGCTGCTTTGATAGTCTGAGGTGCATTTCTGTATGCGACCAAAATGCTGGGTCC 53218
QY 992 TGTTTTCATATCCATCTGTTTACCAAGTATCTTTTCTAGAGAAATTAAGATCATTTAGT 1051
Db 53217 TACTTACATATCCAGTCTGTTAGTCCATGCTCTTTTATTAGAGATTTGAG-----TC 53166
QY 1052 CATTGATGTTGAGAAATTAATCAATGACGAGTGTGTTGGATCTTTGTATCTTGCACTGT 1111
Db 53165 CTTTTCATGTTAAGAGATATTAAAGGAAAGTGAATGTTTACTTCTCTATTTTGTGTTA 53106
QY 1112 GAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1171
Db 53105 GAGTGGAAATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 53063
QY 1172 TGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1231
Db 53062 -----T 53062
QY 1232 TGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1291
Db 53061 GAAAGAAGATACCTTTCTGCTTTTCTAGGGTGTAGTTTCCCTCCTTGTGTTGGAGTTT 53002
QY 1292 TCTCTAGCTCTCT--TAGGTCGTCATTTGAAGATAGATATCTTTACATCTGATTTTA 1349
Db 53001 TTGCTGTTTATCTTTTGAAGGCTGGAATTTGGAAGAATATTGTTAAATTTGTTTGTG 52942
QY 1350 TCTTGAATGCTCTTTCTCCAACTAATTTGTCACAGAAAGTTTCTTAAGTGCAGTAGT 1409
Db 52941 TCATGGAATATATTGTTTTCACGTCATGTTAAATTGAGAGTTTCTGCTGGGTATAGTAGC 52882
QY 1410 CTGGCTGACATCTGATGCTCTTGGAGTCTGTAGACATCTGTGCGAGGCTCTTAC 1469
Db 52881 CTGGACTGATATTGTTGTTCTCTCAGGGTCTGTATGAGATCTGCTAAGATCTCTTAGCT 52822
QY 1470 TTTTGAAGTTCTATTGGAAGTCTAGGTGTAATTTCTAATACATCTGCTCTTATATGTTAA 1529
Db 52821 TTCACAGTCTCT-TGTGAGAGTCTGTTTAAATTTCTGAGAGTCTGCTCTTATATGTTAC 52763
QY 1530 TTGGTCTTTTTCCTTGGCATCTTTTAAATTTCTTTTCTTTGTTCTATACCTTTTGTAGTAT 1589
Db 52762 TT-GACCTTTTCCCTTACTGCTTTTAAATATATCTTTCTTTGTTTGTGCAATTTGTGTTT 52704
```

```
QY 1590 TGATTATTATGCACTGTGGGAGTTTCTTTTCCGTTCCAATCTATTTTGGTGTGTTGTATG 1649
Db 52703 TGATTATTATGATGCGGAGGAATTTCTTTTCTGGTCCAGTCTATTGAGATCTGTAGG 52644
QY 1650 CTTCTTGTACCTTGATAGGACATCTCTTCTCAAGGTTAGGAAATTTTCTTTT---TTG 1705
Db 52643 CTTCTTGTATGTTTCATGGGCATCTCTTTCTTTAGGTTAGGGAAGTTTCTTCTATACATTT 52584
QY 1706 GTTTTCTTGAATAATTTTCCCTGCTTTTGACCTGCTTCTTCCCTCC-----TCT 1758
Db 52583 TTTTGTGTAAGATATTTTACTGACCTTTTAAGTGTGTAATTTTCTATACCTTAATATCC 52524
QY 1759 ATTCTCTTGTGTTTTCATAGTGTCTCTGGCTTCTGGAATGTTTATGCTTGAATTTT 1818
Db 52523 TTAGATTGTTCTTCCATTTGCTCTGATTTCTTGGATGTTTGGTTAGTAGGAGCTTTT 52464
QY 1819 TAGACTTAACATTTCTTTTGACCAAGGATCATTTCTTCTATCTGTCTTCACTGCTG 1878
Db 52463 TGCATGTTGTTTCTTTTCTGACTGTTCTGTTGATGTTTATGTTGTTTCTTCTGACCTG 52404
QY 1879 AGATTCTCTCTCTATCTCTGTTATTTCTGTCAGTGAGGCTTGTCTGAGGTTCTCTG-TT 1937
Db 52403 AGATTCTCTCTCTATCTCTGTTATTTCTGTTGGATGCTTGCATCTATAAATCTCTGATC 52344
QY 1938 GGTTTCTTAATTTTTCATTTCCAGATTTCTTTCAGTTTGGGTTTGTGTTTATTAATCTA 1997
Db 52343 TCTTCTAGGTTTCTTAACTCCAGGCTGTCTCTTTTGTATCTTCTTTATGTTTCTA 52284
QY 1998 TTTCCACTTTCAGTCTCGAATGTTTCTACTCATTTTCTCT---CCAGATTTTACATTTT 2054
Db 52283 TTTCTCTTTTATGATCTTGATGTTTGTTCATTTCTTCCCTTGGCTATTGATGTTGTTT 52224
QY 2055 CATAGGTTTCTTTAATGGAATTTATTCATTTCTCTCTTC 2091
Db 52223 CCTGTCACTTCTTTAAGGATTTTGTGTTGTGTTTC 52187
```

RESULT 35

```
US-10-322-281-495
; Sequence 495, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 52945201000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 495
; LENGTH: 176594
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(176594)
; OTHER INFORMATION: n = A,T,C or G
US-10-322-281-495
```

Query Match

```
Best Local Similarity 24.5%; Score 558.6; DB 1; Length 176594;
Matches 1149; Conservative 0; Mismatches 559; Indels 130; Gaps 19;
QY 292 TCTATTTCTTGATTTCTATCTTGGCTCATTTTAACTCAGTAGTAGTGTGTTGGTTTCC 351
Db 129540 TTTCTTTCTTTTATTCCTTCTTCCCAAGGATATCATTTGAGAAGAGTGTGTTCCGTTTC 129599
QY 352 ATAGTTTGTGAAGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 411
Db 129600 ACGTGAAGTGGCTTTCCCAATATCTATGTTGTTATGAAGATC--AGCCTTAGTCCAT 129656
QY 412 GGTGGTCAGATAGGACATAGATGATTAATTTCAATGTGCTTTTATCTGTCGAGACTTGCTT 471
```

Db 129657 GGTGATCTGATAGGATCGGACAAATTTCAATATATTTTGTATCTGTGTGGGCGCTGTT 129716
QY 472 TGTGTTTGAATATGTAATCAATTTTGGAGA--GTTTCATAGGCTGCTGACAGAGGTAC 529
Db 129717 TGTGACCAATATATAGTCAATTTTGGAGAAGGTACCATGAGGTGCTGAGAAGATGTAT 129776
QY 530 AGTCCTTTGTGTTTGG-TGAAATAGTCTCTAAATATCT-CTAGGTCCCACTTGGTTTATGA 587
Db 129777 ATCCCTTTGGTTTGGATAAAATGTTCTGTAGATATCTGTTCAATCTATTTGTTTCAAA 129836
QY 588 CATCAGTGTAGCTCCAGCAATTTCTGTGTTTGGTTTGTGTGATGACCTAACTGTTGGA 647
Db 129837 CTTCCTGTTAGTTTCACTGTGTCCTGTTTAG-TTTCGTGTTTCCATGATCTGTTCATTGGT 129895
QY 648 GAGATGGGGTATTGAAGTAGGCCACTATCTGTGTGAGGT-CAATATGTGATTTTACG 706
Db 129896 GAGAGTGGGGTGTGAATCTCCACTATATTTGTGTGAGATGCAATGTGTGTTGAAC 129955
QY 707 TGTAGCTGTCTGTTTGTGAACTTGGGTGACATGTGTTGGTGATAGACATAAGA 766
Db 129956 TTTACTAAAGTTTCTTTTAAATAATGGCTGCCCTGCAATTTGGAGCATAGATATTCAGA 130015
QY 767 ATTGCAATCTCTCTTGGTGGATTTT-CCTTTGATGCCCTATGTAGTATTTCTTCCAAATCT 825
Db 130016 ATTGAGAGTCTCTTGGAGATTTTACCTTTGATGAGTATGAAGTGTCCCTCTCTGTCT 130075
QY 826 CATCTGCTTAGTTTGGGTTTAAGTCTA-TTAGTCCAGATATTAATAAGTCTGTATCGGCT 884
Db 130076 TTTTGTGATAACTTTTGGTGGAGTCCGATTTTATTTCAATATTAGAAGGCTACTCCACCT 130135
QY 885 TGTCTCTTAGGCGCATTTGCTTAGATA-TCTTTTCCATCTCTTTTACTCTAAGGTGATGT 943
Db 130136 TGTTCCTCAGGCCAATTTCTTGGAAAAATGTTTCCAGCCCTTCAATCTGAGGTAGTGT 130195
QY 944 CTATCCAT---GGTAGGTGTCTTTTGGATGACAGCATAGGATGGATCTTTGTTTTCAT 1000
Db 130196 CTGTCTTTTCCCTCGGATGGGTTTCTGTGAAGCAGCAAAATGTTGGGCTCTGTTGTGT 130255
QY 1001 ATCCATCTGTACCAGTATCTTTTCTAGAGAAATTAAGATCAATGAGTCATGTATGT 1060
Db 130256 AGCCAGTCTGTAGTCTATGTCTTTTATTTGGGAATTGAG-----TCCATTTGATAT 130307
QY 1061 TGAGAAATATCAATGAGCAGTGTGTTGTGATTTCTTGTATCTTGCACCTTGTGAAGTGTGT 1120
Db 130308 TAAGAGATATTAAGAAAGTAATTTGTTGCTCTCTATTTATTTTGTGTTAAAGTTGCCA 130367
QY 1121 GT 1180
Db 130368 TTCTGTCTTGTGGCTGTCTCTTTTAGGTTTG----- 130400
QY 1181 GTGTCTGTGTCTGTGTGTGTGTCTCTCCCTCTTTTGATTTTGGCCTGGAATAT 1240
Db 130401 -----TTGAAGGAT 130409
QY 1241 TPAATATTCAATTTCTTGAATGGGTAAACATCTTTAGATTTGAAGTTTTCCTCTAGC 1300
Db 130410 TACTTCTGTCTTTCTAGGGTGTCTTCCATCTTTATTTGTTTTTTTTTTTTTCT 130469
QY 1301 CTCTTT-----TAGTCTGCATTTGAAGATAGATTTCTTACATCTGATTTATCT 1352
Db 130470 GTTGTATCTCTTGAAGGGCTGGATTCATCGAAAGATAAATGTGCAATTTGGTTTGTCTG 130529
QY 1353 TAGAATGTCTTTCTTTCTCCAACTATGTGACAGAAAGTTTTTCTAAGTGCAGTAGTCTG 1412
Db 130530 TGGATATCTTTGGTTTCTCCATCTATGTATTTGAAGTTTGGCTGGGTATAGTACCTG 130589
QY 1413 GCCTGACATCTGTAGTCTCTTGGAGTCTGTAGCAGCATCTGTGCGAGGCCCTTCTTACATTT 1472
Db 130590 GGCTGGCATTTGTGTCTCTTGT 130649
QY 1473 TGAGTTTCTATTGGAAGAGTCAGGTGTAATTTCAATACATCTGCTTTATATGTTAATG 1532

Db 130650 ATAGTCTCTGGT-GAAAAGTCTGGTGTGATTTCTAATAGGCCCTGCCCTTATATGTTACTTG 130708
QY 1533 GTCTTTTTCCTTGCATCTTTTAAATATATCTTTCTTTGTTCTATATCTTTTAGTATTTGA 1592
Db 130709 ACCTTTTTTCCT-TTATTTGCTTTTAAATATCTATCTTTTAT-TGGTGTGTTTTTGTCTGA 130766
QY 1593 TTATTTATGCACTGTGGGGAGTTTCTTTTCCGGTCCCAATCTATTTGGTGTGTTTGTATGCTT 1652
Db 130767 TTATTTATGTCAGGAGGAATTTCTTTCTAGTCCAGTCTATTTGGAGTCTGTAGGCTT 130826
QY 1653 CTGTGACCTTGATAGGATCTTTCTTCARAGGTAGGAAATTTTCTTTTGGTGTGTTTCT 1712
Db 130827 CTGTGATGTTTCATGGGCATCTCTTTCTTTAGGATGGGAAGTTTCTTCTATAAATTTGT 130886
QY 1713 TGAATAATTTTCCCTGCTTTTGCACCTG-----CCTTCTTCCCCTTCTCT 1758
Db 130887 TGAAGATATTTGCTGCCCTTTAAGTTGAAATCTTTCTTCTGCTCTACTCTATATAT 130946
QY 1759 ATTCTTTGGTTTTTGGATAGTCTCTGCTTCTCTGGATGTTTATGCTGCTGATTTATTT 1818
Db 130947 GTAGGTTTTGTCTTCTCATTTGTCTCTAAATTTCTTAGATGTTTTCAGTCCAGATCTTT 131006
QY 1819 TAGACTTAAATTTTCTTTGACCAAGGTATCCATTTCTTCTATCTGTCTTCACTGCCCTG 1878
Db 131007 TGCATTTGCAATTTCTTTGATGTTGTGCGGATGTTCTATGTAATCTTCTGACCTG 131066
QY 1879 AGATTCCTCTCTCTCTCTCTGTTATCTGTGAGTGGGCTTGTCTGAGGTCTCTG--T 1936
Db 131067 AGATTCCTCTCTCTCTCTGTTATCTGTGCTAATGCTGATCTATGTTCTAGATT 131126
QY 1937 TGGGTTCTTAATTTTCTATTTCCAGATTTCTTCAATTTGGGTTTTGTTTATTAATCT 1996
Db 131127 TTTTCCCTAGGTTTCTATCTCCAGGTTGCTCACTTTGGGTTTTCTTATGTGTCT 131186
QY 1997 ATTTCCACTTTTCCAGTCTCTGAAATGTTTACTCAITTT--TCCTCCAGTATTTACATTTT 2054
Db 131187 ACTTCACTTTTAGTCTTGTATGATTTTGTTCATTTCCACCACCTGTTTGGTGTGTTT 131246
QY 2055 CATAGTCTTTTATGGAATTTATTCATTTCTCTCTCA 2092
Db 131247 TCCTGTTTTTCTTAAGGATTTGTACTGTACCTCTTTA 131284

RESULT 36

US-10-175-523-75
; Sequence 75, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, David
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/10795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1